

From: Chan, Christina
Sent: Thursday, May 08, 2003 3:00 PM
To: Yu, Misook; STIC-Biotech/ChemLib
Subject: RE: 09/851422 == erroneous file processed

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Yu, Misook
Sent: Thursday, May 08, 2003 2:56 PM
To: Chan, Christina
Subject: FW: 09/851422 == erroneous file processed

The case is 8 month amended and previous search was improperly done. Please approve rush search for SEQ ID NO: 1, 2, and 8.

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

-----Original Message-----

From: Corrigan, Anne-Marie
Sent: Thursday, May 08, 2003 1:40 PM
To: Yu, Misook
Cc: Mosher, Mary
Subject: 09/851422 == erroneous file processed

Exr. Yu:

Unfortunately, STIC processed the wrong file under the above serial number (listed in search system as US09851422B). I removed the erroneous files from the search system; now, the files corresponding to the "ENTERED" raw sequence listing (09/851422A) in your file wrapper can be searched. If you'd like, I'd be happy to contact the searcher who produced your search.

I apologize for the error.

RECEIVED
MAY - 8 2003
BIOT/CHEN L. J. J. J.
(STIC)

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 5/8
Date Completed: 5/12
Searcher Prep/Review: 15
Clerical: _____
Online time: 15

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5 4578
Copyright (c) 1995 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:07:45 : Search time 52.2353 Seconds
(without alignments)
94.366 Million cell updates/sec

Title: US-09-851-422A-1

Sequence: 1 GATTCCTGXXXXXXVDPGIXKXXXXXLIQIEDKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Database : Listing first 45 summaries

- 1: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
- 2: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
- 3: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
- 4: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
- 5: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
- 6: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
- 7: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
- 8: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
- 9: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
- 10: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
- 11: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
- 12: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
- 13: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
- 14: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
- 15: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
- 16: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
- 17: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
- 18: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
- 19: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
- 20: /SID52/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
- 21: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
- 22: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
- 23: /SID52/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	102	73.9	25	AA021730	Amoebapore helix 3
2	102	73.9	25	AA021730	Procytoxin 1
3	102	73.9	25	AA021730	Procytoxin 2
4	102	73.9	25	AA021730	Procytoxin 3
5	102	73.9	25	AA021730	Procytoxin 4
6	102	73.9	25	AA021730	Procytoxin 5
7	102	73.9	25	AA021730	Procytoxin 6
8	102	73.9	25	AA021730	Procytoxin 7
9	102	73.9	25	AA021730	Procytoxin 8
10	101.3	73.8	27	AA021730	Procytoxin 9

11	97.5	70.7	24	AA021741	Procytoxin 10
12	97.5	70.7	24	AA021741	Procytoxin 11
13	97.5	70.7	24	AA021741	Procytoxin 12
14	97.5	70.7	24	AA021741	Procytoxin 13
15	97.5	70.7	24	AA021741	Procytoxin 14
16	97.5	70.7	24	AA021741	Procytoxin 15
17	97.5	70.7	24	AA021741	Procytoxin 16
18	97.5	70.7	24	AA021741	Procytoxin 17
19	97.5	70.7	24	AA021741	Procytoxin 18
20	97.5	70.7	24	AA021741	Procytoxin 19
21	97.5	70.7	24	AA021741	Procytoxin 20
22	97.5	70.7	24	AA021741	Procytoxin 21
23	97.5	70.7	24	AA021741	Procytoxin 22
24	97.5	70.7	24	AA021741	Procytoxin 23
25	97.5	70.7	24	AA021741	Procytoxin 24
26	97.5	70.7	24	AA021741	Procytoxin 25
27	97.5	70.7	24	AA021741	Procytoxin 26
28	97.5	70.7	24	AA021741	Procytoxin 27
29	97.5	70.7	24	AA021741	Procytoxin 28
30	97.5	70.7	24	AA021741	Procytoxin 29
31	97.5	70.7	24	AA021741	Procytoxin 30
32	97.5	70.7	24	AA021741	Procytoxin 31
33	97.5	70.7	24	AA021741	Procytoxin 32
34	97.5	70.7	24	AA021741	Procytoxin 33
35	97.5	70.7	24	AA021741	Procytoxin 34
36	97.5	70.7	24	AA021741	Procytoxin 35
37	97.5	70.7	24	AA021741	Procytoxin 36
38	97.5	70.7	24	AA021741	Procytoxin 37
39	97.5	70.7	24	AA021741	Procytoxin 38
40	97.5	70.7	24	AA021741	Procytoxin 39
41	97.5	70.7	24	AA021741	Procytoxin 40
42	97.5	70.7	24	AA021741	Procytoxin 41
43	97.5	70.7	24	AA021741	Procytoxin 42
44	97.5	70.7	24	AA021741	Procytoxin 43
45	97.5	70.7	24	AA021741	Procytoxin 44

ALIGNMENTS

RESULT 1
AA021730 standard: peptide: 25 AA.
AA021730:
13-SEP-2002 (first entry)
Amoebapore helix 3 cytolitic peptide.
Cytotoxic, cytosolic; procytoxin; protease; cancer; ovary; prostate; breast; skin; lung; pancreas; amoebapore helix 3.
Entamoeba histolytica.
US2002045736-A1.
18-APR-2002.
27-AUG-2001: 2001US-0938623.
09-MAY-2001: 2001US-0851422.
(YUXX) YU X.
(WGNH) WAGNER T E.
Yu X, Wagner TE.
WPI: 2002-507251/54.
A new procytoxin useful in the treatment of cancer of e.g. prostate, ovary, breast, or skin, has a cytolitic peptide bound to an inactivator via a peptide bond cleavable by a specific protease.

PR	09-MAY-2000:	2000US-203063P.
PR	16-JUN-2000:	2000US-212042P.
XX	(GRR-) GREENVILLE HOSPITAL SYSTEM.	
XX		
PI	Yu X., Wagner TE:	
DR	WPI: 2002-154423/20.	
PT	A proteotoxin comprising a cytotoxic peptide with at least one lysine residue which is converted into a cytotoxin in a target cell is used for treating cancer, particularly of the prostate, skin, ovary or lung	
PS	Dislosure: Page 7, 33pp: English.	
CC	The present invention relates to a proteotoxin, comprising a cytotoxic peptide with at least one lysine residue bound by a peptide bond to at least one amino acid via the epsilon-amino group of the lysine residue.	
CC	One proteotoxin is used to treat cancer, particularly of the prostate, skin, ovary or lung.	
CC	Another proteotoxin is used to treat cancer, causing cell lysis and death (apoptosis) typically disrupt cell membrane, causing cell lysis and death	
CC	procytotoxin, amebapore helix 3.	
SQ	Sequence 25 AA:	
Overy Match	73.8% Score 102; DB 23;	
Best Local Similarity	75.8%; Pred. No. 5e-12;	
Matches 25:	Conservative 0; Mismatches 0; Indels 8; Gaps 2	
Oy	1 GTATLCRXXXXYVDFEGRKXXXLLQLIIDEKD 33 1 GTATLCRK---VLDFSIKK---LIIIDLDK 25	
Db		
RESULT 4		
AAE18197	AAE18197 standard: peptide: 25 AA.	
AC	AAE18197:	
XX	07-MAY-2002 (first entry)	
DE	Procytotoxin #1 for treating cancer.	
KM	Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;	
KW	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.	
OS	Unidentified.	
Key		
Modified-site	Location/Qualifiers	
FT	/note= "lys(R), where R is the unmodified epsilon-amino group of the adjacent Lys, or is [epsilonion-gamma]-Glu,	
FT	[epsilonion-gamma]-Glu-[alpha-gamma]-(Glu)1-3,	
FT	[epsilonion-gamma]-(the)1-3, [epsilonion-alpha]-(tyr)1-3,	
FT	[epsilonion-alpha]-(trp)1-3, [epsilonion-alpha]-(lys)1-3 or	
FT	[epsilonion-alpha]-(arg)1-3, [epsilonion-alpha]-(asp)1-3, and	
FT	represents a peptide bond between the epsilon-amino group of lysine and the gamma carboxyl group of the adjacent glutamate, [alpha-gamma] represents a peptide bond between the alpha amino group of the second glutamate and the gamma carboxyl group of the first glutamate, [epsilonion-gamma] represents a peptide bond between the epsilon-amino group of lysine and the alpha carboxyl group of the adjacent Glu, or is [epsilonion-gamma]-Glu,	
FT	[epsilonion-gamma]-Glu-[alpha-gamma]-(Glu)1-3,	
FT	or the first via conventional peptide bonds"	
FT	17	
FT	/note= "lys(R), where R is the unmodified epsilon-amino group of the adjacent Lys, or is [epsilonion-gamma]-Glu,	
FT	[epsilonion-gamma]-Glu-[alpha-gamma]-(Glu)1-3,	
FT		

[illegible]

[illegible]

XX	XX	Unidentified.	
XX	XX	Key	Location/Qualifiers
XX	XX	Modified-site	/note= "The epsilon amino group of lys 1s peptide bonded to the alpha carboxyl group of phe"
FT	FT	Modified-site	25
FT	FT	Modified-site	/note= "The epsilon amino group of lys 1s peptide bonded to the alpha carboxyl group of phe"
FT	FT	Modified-site	26
FT	FT	Modified-site	/note= "The alpha carboxyl group of phe 1s peptide bonded to the epsilon amino group of lys"
XX	XX	US2002045736-A1.	
XX	XX	18-APR-2001.	
PD	PD	27-AUG-2001: 2001US-0938623.	
XX	XX	09-MAY-2001: 2001US-0851422.	
PA	PA	(YUXX) YU X.	
XX	XX	(WAGN) WAGNER T E.	
XX	XX	yu X, wagner TE;	
XX	XX	WPI: 2002-507251/54.	
XX	XX	A new procytotoxin useful in the treatment of cancer of e.g. prostate, ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator via a peptide bond cleavable by a specific protease .	
XX	XX	Disclosure: Page 7: 21pp: English.	
XX	XX	The invention relates to a procytotoxin comprising a cytotoxic peptide bound to an inactivator via a peptide bond, where the peptide bond is susceptible to cleavage by a targeting specific protease. The procytotoxin is used to treat cancer, particularly of the prostate, ovary, breast, skin, lung or pancreas. This sequence represents a procytotoxin cytolytic peptide sequence relating to the invention.	
XX	XX	Sequence 26 AA:	
XX	XX	Query Match 73.9%: score 102: DB 23: length 26:	
XX	XX	Best Local Similarity 75.8%: Pred. No. 5.7e-12:	
XX	XX	Matches 25: Conservative 0: Mismatches 0: Indels 8: Caps 2	
XX	XX	1 GFATLCTKXXXXXVDFGIDKXXXXLLOLIDBK 33	
XX	XX	1 GFATLCTK-----VDFGIDK-----LLOLIDBK 25	
XX	XX	RESULT 7	
XX	XX	AAEL8200	
XX	XX	ID AAEL8200 standard: peptide: 26 AA.	
XX	XX	AAEL8200:	
XX	XX	07-MAY-2002 (first entry)	
XX	XX	Proloxin #1 for treating cancer.	
XX	XX	Procytotoxin: cytotoxic peptide: epsilon-amino group: cancer: cytostatic:	
XX	XX	prostate: ovary: lung: skin: channel-forming peptide: apoptosis: therapy.	
XX	XX	Unidentified.	
XX	XX	Key	Location/Qualifiers
XX	XX	Modified-site	25
XX	XX	Modified-site	/note= "lys-[epsilon-alpha]: where [epsilon-alpha] represents a peptide bond between the epsilon amino group of lysine and the alpha carboxyl group of the

adjacent phenylalanine"

W0200185777-A2.

15-NOV-2001.

09-MAY-2001: 2001WO-US40690.

09-MAY-2000: 2000US-203063P.

16-JUN-2000: 2000US-212042P.

(GHEE-) GREENVILLE HOSPITAL SYSTEM.

Yu X, Wagner TE;

WPI: 2002-154423/20.

A procytotoxin comprising a cytotoxic peptide with at least one lysine residue which is converted into a cyclooxin in a target cell is used for treating cancer, particularly of the prostate, skin, ovary or lung

Disclosure: Page 12: 33pp: English.

The present invention relates to a procytotoxin, comprising a cytotoxic peptide with at least one lysine residue bound by a peptide bond to at least one glutamic acid residue, which is converted into a cyclooxin of the prostate, ovary, lung or skin. Cytolytic peptide also known as channel-forming peptide, typically disrupt cell membrane, causing cell lysis and death (apoptosis) upon contact. The present sequence is protoxin.

Sequence 26 AA:

Query Match 73.9% Score 102: DB 23: Length 26:
Best Local Similarity 75.8%: Pred. No. 5,7e-12:
Matches 25: Conservative 0: Mismatches 0: Indels 8: Gaps 2:

QY 1 GFATICTGXXXXVDFGIDKXXXXLIQLIEK 33
DB 1 GFATICTG----VDFGIDK----LIQLIEK 25

RESULT 8

AA021736

AA021736 standard: Peptide: 27 AA.

AA021736:

13-SEP-2002 (first entry)

Procytotoxin cytolytic peptide #1.

Cytotoxic; cytosstatic; procytotoxin; inactivator; protease; cancer; ovary; prostate; breast; skin; lung; pancreas; cytolitic.

Unidentified.

Key Location/Qualifiers

Modified-site 25 /note- "The epsilon amino group of lys is peptide bonded to the gamma carboxyl group of Glu"

Modified-site 26 /note- "The gamma carboxyl group of this Glu is peptide bonded to the epsilon amino group of Lys. The alpha amino group of another Glu"

Modified-site 27 /note- "The gamma carboxyl group of this Glu is peptide bonded to the alpha amino group of another Glu"

US2002045736-A1.

18-APR-2002.

27-AUG-2001: 2001US-0938623.

09-MAY-2001: 2001US-0851422.

(YUXK/) YU X.

(MAGN/) MAGNER T E.

Yu X, Wagner TE;

WPI: 2002-507251/54.

A new procytotoxin useful in the treatment of cancer of e.g. prostate, PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator via a peptide bond cleavable by a specific protease

Disclosure: Page 7: 21pp: English.

The invention relates to a procytotoxin comprising a cytotoxic peptide bound to an inactivator via a peptide bond, where the peptide bond is susceptible to cleavage by a targeting specific protease. The procytotoxin is used to treat cancer, particularly of the prostate, ovary, breast, skin, lung or pancreas. This sequence represents a procytotoxin cytolitic peptide sequence relating to the invention.

Sequence 27 AA:

Query Match 73.9% Score 102: DB 23: Length 27:
Best Local Similarity 75.8%: Pred. No. 5,9e-12:
Matches 25: Conservative 0: Mismatches 0: Indels 8: Gaps 2:

QY 1 GFATICTGXXXXVDFGIDKXXXXLIQLIEK 33
DB 1 GFATICTG----VDFGIDK----LIQLIEK 25

RESULT 9

AAE18199

AAE18199 standard: peptide: 27 AA.

AAE18199:

07-MAY-2002 (first entry)

Procytotoxin #3 for treating cancer.

Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytosstatic; prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.

Unidentified.

Key Location/Qualifiers

Modified-site 25 /note- "Lys-[epsilon-[gamma]]: where [epsilon-[gamma]] represents a peptide bond between the epsilon amino group of lysine and the gamma carboxyl group of the adjacent glutamate"

Modified-site 26 /note- "Glu-[alpha-[gamma]]: where [alpha-[gamma]] represents a peptide bond between the alpha amino group of the first glutamate and the gamma carboxyl group of the second glutamate"

W0200185777-A2.

15-NOV-2001.

09-MAY-2001: 2001WO-US40690.

09-MAY-2000: 2000US-203063P.

16-JUN-2000: 2000US-212042P.

[illegible][illegible]

Query Match 70.7% Score 97.5: DB 23: Length 24:
Best Local Similarity 72.7%: Pred. No. 3, 7e-11:
Matches 24: Conservative 0: Mismatches 0: Indels 9: Gaps 2:
OY 1 GFIATCTKXXXXXVDFGIDKXXXXXLIQLEDK 33
DB 1 GFIATCTK-----VLDGIDK-----IQLLEDK 24
RESULT 12
AAE18203
AAE18203 standard: peptide: 24 AA.
AC AAE18203:
XX 07-MAY-2002 (first entry)
XX Cytolytic peptide for treating cancer.
DE Procytotoxic peptide: epsilon-amino group; cancer: cytostatic;
KM prostate; ovary: lung; skin: channel-forming peptide; apoptosis; therapy.
XX Unidentified.
XX MO200185777-A2.
XX 15-NOV-2001.
PE 09-MAY-2001: 2001MO-US40690.
PR 09-MAY-2000: 2000US-203063P.
PR 16-JUN-2000: 2000US-212042P.
XX (GREF-) GREENVILLE HOSPITAL SYSTEM.
XX Yu X, Wagner TE;
PI WPI: 2002-154423/20.
DR A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX Example 1: Page 20: 33pp: English.
PS The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is cytolytic peptide.
XX
XX Sequence 24 AA:
Query Match 70.7% Score 97.5: DB 23: Length 24:
Best Local Similarity 72.7%: Pred. No. 3, 7e-11:
Matches 24: Conservative 0: Mismatches 0: Indels 9: Gaps 2:
OY 1 GFIATCTKXXXXXVDFGIDKXXXXXLIQLEDK 33
DB 1 GFIATCTK-----VLDGIDK-----IQLLEDK 24
RESULT 13
AAE18204
AAE18204 standard: peptide: 26 AA.
AC AAE18204:
XX 07-MAY-2002 (first entry)
XX

DE Procytolytic peptide for treating cancer.
XX
XX Procytotoxin: cytotoxic peptide; epsilon-amino group; cancer: cytostatic;
KM prostate; ovary: lung; skin: channel-forming peptide; apoptosis; therapy.
XX Unidentified.
XX
XX Key
FH Modified-site 23
FT /note- "Addition of two gamma glutamate linked side-chain
FT glutamate acid residue to the epsilon amino group of the
FM C-terminal lysine"
XX
XX Location/Qualifiers
XX MO200185777-A2.
XX 15-NOV-2001.
PD 09-MAY-2001: 2001MO-US40690.
XX 09-MAY-2000: 2000US-203063P.
XX 16-JUN-2000: 2000US-212042P.
XX (GREF-) GREENVILLE HOSPITAL SYSTEM.
XX Yu X, Wagner TE;
PI WPI: 2002-154423/20.
DR A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX Example 1: Page 20: 33pp: English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procytolytic peptide.
XX
XX Sequence 26 AA:
Query Match 70.7% Score 97.5: DB 23: Length 26:
Best Local Similarity 72.7%: Pred. No. 4e-11:
Matches 24: Conservative 0: Mismatches 0: Indels 9: Gaps 2:
OY 1 GFIATCTKXXXXXVDFGIDKXXXXXLIQLEDK 33
DB 1 GFIATCTK-----VLDGIDK-----IQLLEDK 24
RESULT 14
AAE18205
AAE18205 standard: peptide: 26 AA.
AC AAE18205:
XX 07-MAY-2002 (first entry)
XX Procytotoxin #5 for treating cancer.
DE Procytotoxin: cytotoxic peptide; epsilon-amino group; cancer: cytostatic;
KM prostate; ovary: lung; skin: channel-forming peptide; apoptosis; therapy.
XX Unidentified.
XX
XX Key
FH Modified-site 24, 26
FT /note- "Lys((epsilon-gamma)-Glu-(alpha-gamma)-Glu);
FT where [epsilon-gamma] represents a peptide bond between
XX

FT the epsilon amino group of lysine and the gamma carboxyl
FT group of adjacent glutamate, [alpha-gamma] represents a
FT peptide bond between alpha amino group of glutamate and
FT gamma carboxyl group of adjacent glutamate
PN MO200185777-A2.
PD 15-NOV-2001.
XX 09-MAY-2001: 2001MO-US040690.
PF 09-MAY-2000: 2000US-203063P.
PR 16-JUN-2000: 2000US-212042P.
XX (GREE-) GREENVILLE HOSPITAL SYSTEM.
PA
PL Yu X. Wagner TE;
PI WPI: 2002-154437/20.
XX
XX A procyctoxin comprising a cytotoxic peptide with at least one lysine
FT residue which is converted into a cycotoxin in a target cell is used
FT for treating cancer, particularly of the prostate, skin, ovary or lung
PF
XX
XX Claim 8; Page 24; 33pp; English.
XX
XX The present invention relates to a procyctoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procyctoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cycolytic peptide also known as channel-forming
CC peptide is known to cause cell lysis, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procyctoxin.
SQ Sequence 26 AA:
Query Match 70.7%; Score 97.5; DB 23; Length 26;
Best Local Similarity 72.7%; Pred. No. 4e-11; 0; Mismatches 9; Gaps 2;
Matches 24; Conservative 0; Indels 9;
C1 1 GFIATLCTKXXXXXVLDGIDKXXXXXITQLIEEM 33
D1 1 GFIATLCTK-----VLDGIDK-----ITQLIEEM 24
DB 1 GFIATLCTK-----VLDGIDK-----ITQLIEEM 24
RESULT 15
ACAC1353
ID AAC41753 standard: Proctolin; 463 AA.
XX AAC41753:
AC AAC41753:
XX
XX 18-OCT-2000 (first entry)
DT
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51987.
XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hydrolisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX EPI03405-A2.
XX
XX 06-SEP-2000.
PD
XX
PF 25-FEB-2000: 2000EP-0301439.
XX
XX 25-FEB-1999: 9905-0121825.
PR 05-MAR-1999: 9905-0123180.
PR 09-MAR-1999: 9905-0123548.
PR 09-MAR-1999: 9905-0123548.
PR 25-MAR-1999: 9905-0126264.
PR 25-MAR-1999: 9905-0126264.
PR 29-MAR-1999: 9905-0126765.
PR 01-APR-1999: 9905-0127452.
PR 06-APR-1999: 9905-0128234.
PR 08-APR-1999: 9905-0128234.
PR 08-APR-1999: 9905-0128244.
PR 19-APR-1999: 9905-0130077.
PR 21-APR-1999: 9905-0130449.
PR 23-APR-1999: 9905-0130510.
PR 23-APR-1999: 9905-0130891.
PR 28-APR-1999: 9905-0131449.
PR 30-APR-1999: 9905-0132048.
PR 30-APR-1999: 9905-0134216.
PR 14-MAY-1999: 9905-0134216.
PR 14-MAY-1999: 9905-0134221.
PR 14-MAY-1999: 9905-0134370.
PR 18-MAY-1999: 9905-0134768.
PR 19-MAY-1999: 9905-0134941.
PR 20-MAY-1999: 9905-0135124.
PR 20-MAY-1999: 9905-0135124.
PR 21-MAY-1999: 9905-0135299.
PR 21-MAY-1999: 9905-0135299.
PR 25-MAY-1999: 9905-0136021.
PR 27-MAY-1999: 9905-0136392.
PR 28-MAY-1999: 9905-0136782.
PR 01-JUN-1999: 9905-0137222.
PR 03-JUN-1999: 9905-0137228.
PR 03-JUN-1999: 9905-0137228.
PR 07-JUN-1999: 9905-0137724.
PR 08-JUN-1999: 9905-0138034.
PR 10-JUN-1999: 9905-0138540.
PR 10-JUN-1999: 9905-0138847.
PR 14-JUN-1999: 9905-0139119.
PR 16-JUN-1999: 9905-0139452.
PR 16-JUN-1999: 9905-0139453.
PR 16-JUN-1999: 9905-0139453.
PR 16-JUN-1999: 9905-0139454.
PR 18-JUN-1999: 9905-0139455.
PR 18-JUN-1999: 9905-0139456.
PR 18-JUN-1999: 9905-0139457.
PR 18-JUN-1999: 9905-0139458.
PR 18-JUN-1999: 9905-0139459.
PR 18-JUN-1999: 9905-0139459.
PR 18-JUN-1999: 9905-0139460.
PR 18-JUN-1999: 9905-0139460.
PR 18-JUN-1999: 9905-0139462.
PR 18-JUN-1999: 9905-0139463.
PR 18-JUN-1999: 9905-0139750.
PR 18-JUN-1999: 9905-0139763.
PR 21-JUN-1999: 9905-0139817.
PR 22-JUN-1999: 9905-0139899.
PR 22-JUN-1999: 9905-0140184.
PR 23-JUN-1999: 9905-0140354.
PR 24-JUN-1999: 9905-0140655.
PR 28-JUN-1999: 9905-0140823.
PR 29-JUN-1999: 9905-0140951.
PR 30-JUN-1999: 9905-0141287.
PR 01-JUL-1999: 9905-0141842.
PR 01-JUL-1999: 9905-0142134.
PR 01-JUL-1999: 9905-0142134.
PR 06-JUL-1999: 9905-0143360.
PR 08-JUL-1999: 9905-0143803.
PR 09-JUL-1999: 9905-0143920.
PR 12-JUL-1999: 9905-0142977.
PR 13-JUL-1999: 9905-0143542.
PR 14-JUL-1999: 9905-0143624.
PR 15-JUL-1999: 9905-0144005.
PR 16-JUL-1999: 9905-0144005.
PR 16-JUL-1999: 9905-0144006.

PR 19-JUL-1999: 9905-0144325.
 PR 19-JUL-1999: 9905-0144331.
 PR 19-JUL-1999: 9905-0144332.
 PR 19-JUL-1999: 9905-0144333.
 PR 19-JUL-1999: 9905-0144334.
 PR 19-JUL-1999: 9905-0144335.
 PR 20-JUL-1999: 9905-0144332.
 PR 20-JUL-1999: 9905-0144632.
 PR 20-JUL-1999: 9905-0144634.
 PR 20-JUL-1999: 9905-0144684.
 PR 21-JUL-1999: 9905-0144814.
 PR 21-JUL-1999: 9905-0145086.
 PR 21-JUL-1999: 9905-0145087.
 PR 22-JUL-1999: 9905-0145085.
 PR 22-JUL-1999: 9905-0145087.
 PR 22-JUL-1999: 9905-0145089.
 PR 22-JUL-1999: 9905-0145192.
 PR 23-JUL-1999: 9905-0145145.
 PR 23-JUL-1999: 9905-0145218.
 PR 23-JUL-1999: 9905-0145274.
 PR 23-JUL-1999: 9905-0145313.
 PR 27-JUL-1999: 9905-0145918.
 PR 27-JUL-1999: 9905-0145919.
 PR 28-JUL-1999: 9905-0145951.
 PR 02-AUG-1999: 9905-0146386.
 PR 02-AUG-1999: 9905-0146388.
 PR 03-AUG-1999: 9905-0146388.
 PR 03-AUG-1999: 9905-0146388.
 PR 04-AUG-1999: 9905-0147204.
 PR 04-AUG-1999: 9905-0147302.
 PR 05-AUG-1999: 9905-0147192.
 PR 05-AUG-1999: 9905-0147260.
 PR 06-AUG-1999: 9905-0147303.
 PR 06-AUG-1999: 9905-0147445.
 PR 09-AUG-1999: 9905-0147935.
 PR 09-AUG-1999: 9905-0148171.
 PR 11-AUG-1999: 9905-0148319.
 PR 12-AUG-1999: 9905-0148341.
 PR 13-AUG-1999: 9905-0148654.
 PR 13-AUG-1999: 9905-0148684.
 PR 15-AUG-1999: 9905-0149175.
 PR 17-AUG-1999: 9905-0149436.
 PR 18-AUG-1999: 9905-0149436.
 PR 20-AUG-1999: 9905-0149722.
 PR 20-AUG-1999: 9905-0149723.
 PR 20-AUG-1999: 9905-0149929.
 PR 23-AUG-1999: 9905-0149902.
 PR 23-AUG-1999: 9905-0149920.
 PR 23-AUG-1999: 9905-0150050.
 PR 26-AUG-1999: 9905-0150884.
 PR 27-AUG-1999: 9905-0151065.
 PR 27-AUG-1999: 9905-0151066.
 PR 27-AUG-1999: 9905-0151080.
 PR 30-AUG-1999: 9905-0151303.
 PR 31-AUG-1999: 9905-0151438.
 PR 01-SEP-1999: 9905-0151430.
 PR 07-SEP-1999: 9905-0151750.
 PR 10-SEP-1999: 9905-0153070.
 PR 13-SEP-1999: 9905-0153758.
 PR 15-SEP-1999: 9905-0154018.
 PR 16-SEP-1999: 9905-0154039.
 PR 20-SEP-1999: 9905-0154779.
 PR 22-SEP-1999: 9905-0155119.
 PR 22-SEP-1999: 9905-0155129.
 PR 24-SEP-1999: 9905-0155456.
 PR 28-SEP-1999: 9905-0156458.
 PR 29-SEP-1999: 9905-0156596.
 PR 04-OCT-1999: 9905-0157117.
 PR 05-OCT-1999: 9905-0157753.
 PR 06-OCT-1999: 9905-0157865.
 PR 07-OCT-1999: 9905-0158029.
 PR 08-OCT-1999: 9905-0158232.

PR 12-OCT-1999: 9905-0158369.
 PR 13-OCT-1999: 9905-0159293.
 PR 13-OCT-1999: 9905-0159294.
 PR 13-OCT-1999: 9905-0159295.
 PR 14-OCT-1999: 9905-0159329.
 PR 14-OCT-1999: 9905-0159331.
 PR 14-OCT-1999: 9905-0159637.
 PR 14-OCT-1999: 9905-0159638.
 PR 18-OCT-1999: 9905-0160741.
 PR 21-OCT-1999: 9905-0160767.
 PR 21-OCT-1999: 9905-0160768.
 PR 21-OCT-1999: 9905-0160798.
 PR 21-OCT-1999: 9905-0160814.
 PR 21-OCT-1999: 9905-0160815.
 PR 22-OCT-1999: 9905-0160980.
 PR 22-OCT-1999: 9905-0160981.
 PR 22-OCT-1999: 9905-0160989.
 PR 23-OCT-1999: 9905-0161404.
 PR 23-OCT-1999: 9905-0161406.
 PR 25-OCT-1999: 9905-0161406.
 PR 26-OCT-1999: 9905-0161359.
 PR 26-OCT-1999: 9905-0161360.
 PR 26-OCT-1999: 9905-0161361.
 PR 28-OCT-1999: 9905-0161920.
 PR 28-OCT-1999: 9905-0161920.
 PR 29-OCT-1999: 9905-0161920.
 PR 29-OCT-1999: 9905-0162142.

Query Match 31.9% Score 44: DB 21: Length 463:
 Best Local Similarity 47.4% Pred. No. 11;
 Matches 9: Conservative 3: Mismatches 7: Indels 0: Gaps 0:
 OY 2 F1ATLCTKXXXYIDRGTID 20
 DB 282 F1ATLSCNHAGIIRBGIID 300

Search completed: May 9, 2003, 16:15:34
 Job time : 53.2353 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:13:10 ; Search time 13.4216 Seconds
(without alignments)
81.112 Million cell updates/sec

Title: US-09-851-422a-1

Perfect score: 1 GFIATLCRKXXXXVLDGIDXXXXXLIQLERKXXX 37

Sequence: 1

Scoring table: BIOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database :

1: Issued Patents_AA.*
2: /cgn2.6/pdata/1/aa/5A_COMB pep.*
3: /cgn2.6/pdata/1/aa/6A_COMB pep.*
4: /cgn2.6/pdata/1/aa/6A_COMB pep.*
5: /cgn2.6/pdata/1/aa/6A_COMB pep.*
6: /cgn2.6/pdata/1/aa/6A_COMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	30.4	893	4	US-09-019-160-6
2	41.5	30.1	460	4	US-09-199-517A-132
3	41	29.7	893	2	US-08-706-702-3
4	41	29.7	893	2	US-09-019-160-3
5	41	29.7	893	4	US-09-019-160-7
6	41	29.7	893	4	US-09-019-160-8
7	41	29.7	893	4	US-09-019-160-9
8	41	29.7	893	4	US-09-131-001C-5477
9	40	29.0	144	4	US-09-019-160-9
10	37	26.8	40	2	US-08-807-332B-5
11	37	26.8	40	2	US-09-336-876-5
12	37	26.8	40	2	US-09-336-876-5
13	37	26.8	40	2	US-09-336-876-5
14	37	26.8	40	2	US-09-336-876-5
15	37	26.8	40	2	US-09-336-876-5
16	37	26.8	40	2	US-09-336-876-5
17	37	26.8	40	2	US-09-336-876-5
18	37	26.8	40	2	US-09-336-876-5
19	37	26.8	40	2	US-09-336-876-5
20	37	26.8	40	2	US-09-336-876-5
21	37	26.8	40	2	US-09-336-876-5
22	37	26.8	40	2	US-09-336-876-5
23	37	26.8	40	2	US-09-336-876-5
24	37	26.8	40	2	US-09-336-876-5
25	37	26.8	40	2	US-09-336-876-5
26	37	26.8	40	2	US-09-336-876-5
27	37	26.8	40	2	US-09-336-876-5

28	37	26.8	403	4	US-09-549-098-10	Sequence 10, Appl
29	37	26.8	420	4	US-09-009-271A-4	Sequence 4, Appl
30	37	26.8	420	4	US-08-974-691-8	Sequence 8, Appl
31	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
32	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
33	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
34	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
35	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
36	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
37	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
38	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
39	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
40	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
41	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
42	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
43	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
44	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
45	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-019-160-6
Sequence 6, Application US/09019160
Patent No. 6306588
GROSS INVENTION:
APPLICANT: Chatterjee, Deb K.
APPLICANT: Solus, Joseph
TITLE OF INVENTION: Polymers for Analyzing or Typing Polymorphic
NUMBER OF SEQUENCES: 93
ADDRESS: STEPHEN KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/019,160
APPLICATION DATE: 06-FEB-1998
FILING DATE: 06-FEB-1998
CLASSIFICATION: (To be assigned)
APPLICATION NUMBER:
FILING DATE: 06-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
INVENTOR INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2600
INVENTOR INFORMATION:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-09-019-160-6

? TELEPHONE: 202-371-2600
? TELFAX: 202-371-2540
? INFORMATION FOR SEQ ID NO.: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 893 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: not relevant
? MOLECULE TYPE: protein
DS-09-019-160-7

LENGTH: 144
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4577
Query Match 29.0%; Score 40; DB 4; Length 144;
Best Local Similarity 41.2%; Pred. No. 6.9;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
OY 5 TCTGKXXXVDFGDK 21
DB 96 TCTGKIIITGAKXVGRK 112
RESULT 10
US-08-807-332B-5
Sequence 5, Application US/08807332B
Patent No. 5959074
GPIANTCTXXXXXXXXXX
APPLICANT: Dreyfus, David H.
APPLICANT: Gelfand, Eytan M.
TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
TITLE OF INVENTION: GENE RECOMBINATION
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross
ADDRESS: 700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US/08/807,332B
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Joseph E. Kovarik
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-39
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NOS: 5;
SEQUENCE CHARACTERISTICS:
TYPE: amino acid
STRANDEDNESS: Linear
TOPOLOGY: Linear
MOLECULE TYPE: peptide
US-08-807-332B-5
Query Match 26.8%; Score 37; DB 2; Length 40;
Best Local Similarity 36.8%; Pred. No. 5.5;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
OY 1 GFATLCTXXXXVDFGI 19
DB 17 GHVQTLCTNFRSLADIKGV 35
RESULT 11
US-09-338-876-5
Sequence 5, Application US/09338876
Patent No. 6187584
GENERAL INFORMATION:
APPLICANT: Dreyfus, David H.
APPLICANT: Gelfand, Eytan M.
TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF

TITLE OF INVENTION: GENE RECOMBINATION
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross
ADDRESS: 700 Lincoln St., Suite 3500
CITY: Denver
STATE: CO
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US/09/338,876
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-39
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NOS: 5;
SEQUENCE CHARACTERISTICS:
TYPE: amino acid
STRANDEDNESS: Linear
TOPOLOGY: Linear
MOLECULE TYPE: peptide
US-09-338-876-5
Query Match 26.8%; Score 37; DB 4; Length 40;
Best Local Similarity 36.8%; Pred. No. 5.5;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
OY 1 GFATLCTXXXXVDFGI 19
DB 17 GHVQTLCTNFRSLADIKGV 35
RESULT 12
US-09-134-001C-4599
Sequence 4599, Application US/09134001C
Patent No. 6380370
GPIANTCTXXXXXXXXXX
APPLICANT: Dreyfus, David H.
APPLICANT: Gelfand, Eytan M.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NOS: US 60/064,964
PRIOR FILING DATE: 1997-08-14
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4599
LENGTH: 126
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4599
Query Match 26.8%; Score 37; DB 4; Length 126;
Best Local Similarity 31.0%; Pred. No. 21;
Matches 9; Conservative 4; Mismatches 16; Indels 0; Gaps 0;
OY 1 GFATLCTXXXXVDFGDKXXXXL 29

```

Db          43  G1ATPAPVYLAIALAIABKCOOLYM 71
|||||:      ::  |||:
RESULT 13
US-08-738-944-50
: Sequence 50, Application US/08738944
: Patent No. 5783431
: GENERAL INFORMATION:
: APPLICANT: Peterson, et al.
: TITLE OF INVENTION: METHODS FOR GENERATING AND
: TITLE OF INVENTION: SCREENING NOVEL METABOLIC PATHWAYS
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTED3 version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/738,944
: FILING DATE: 24-APR-1996
: CLASSIFICATION: 536
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US98 08/539,255
: FILING DATE: 24-APR-1996
: ATTORNEY/AGENT INFORMATION:
: ATTORNEY: Cortez, Laura A
: REGISTERED NUMBER: 742
: REFERENCE/DOCKET NUMBER: 8757-007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-790-9090
: TELEFAX: 212-869-8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 50:
: SOURCE CHARACTERISTICS:
: LENGTH: 26 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FEATURE:
: NAME/KEY: CXC-AMN20
: ORIGIN:
: LOCATION: 1..262
: OTHER INFORMATION:
US-08-738-944-50
Query Match          26.8%: Score 37; Length 262;
Best Local Similarity 36.7%: Pred.No. DB:
Matches 11: Conserving 4: Mismatches 13: Indels 2: Gaps 1:
      4  ATCTGXXXXXVDPGIDXXXXXLIQLEIDK 33
      |||
Db    226  ATLRTPKXAYGDF--DTSISLLEPMWIDK 253
|||||
RESULT 14
US-09-263-352-40
: Sequence 40, Application US/09263352
: Patent No. 6242211
: GENERAL INFORMATION:
: APPLICANT: Peterson, T.
: TITLE OF INVENTION: METHODS FOR GENERATING AND
: TITLE OF INVENTION: SCREENING NOVEL METABOLIC
: FILE REFERENCE: 8757-010
: CURRENT APPLICATION NUMBER: US/09/263,352

```

```

1      CURRENT FILING DATE: 1999-03-05
2      EARLIER APPLICATION NUMBER: 08/986,186
3      EARLIER FILING DATE: 1997-12-05
4      NUMBER OF SEQ ID NOS: 41
5      SOURCE: GenBank
6      SEQ ID NO: 40
7      LENGTH: 262
8      TYPE: PRT
9      ORGANISM: Streptomyces coelicolor
10     US-05-263-352-40
11
12     Query Match: 26.8% Score 37, DB 4, Length 262,
13     Seq. Id. 40, Pos. 1, Start 1, End 262,
14     Local Similarity: 26.78%, 47.8%, 50%,
15     Matches 11: Conservative 4, Mismatches 13, Indels 2, Gaps 1,
16
17     DB 226 ATLRTFGKAYDDF-DTISYLLBMMIDK 253
18
19     4 ATLCGTXXXXVDTGGLKXXVLDLEK 33
20       ||| ||| || | |::: ||
21
22     RESULT 15
23     US-08-723-938-3
24     Sequence 3, Application US/08723938
25     Patent No. 5776759
26     GENERAL INFORMATION:
27     APPLICANT: Bandman, Olga
28     ATTORNEY/AGENT INFORMATION:
29     NAME: Billings, Lucy J
30     REGISTRATION NUMBER: 36,749
31     EXPIRATION DATE: 01-01-2005
32     TELECOMMUNICATION INFORMATION:
33     TELEPHONE: 415-855-0555
34     TELEFAX: 415-845-4166
35     TELEX:
36     INFORMATION FOR SEQ ID NO: 3:
37     SOURCE CHARACTERISTICS:
38     LENGTH: 395 amino acids
39     TYPE: amino acid
40     STRANDEDNESS: single
41     TOPOLOGY: linear
42     MOLECULE TYPE: peptide
43     HYPOTHETICAL: NO
44     ANTI-SENSE: NO
45     FRAGMENT TYPE: N-terminal
46     ORIGINAL SOURCE:
47     IMMEDIATE SOURCE:
48     CLONE: 312099
49     STRAIN: J1000702
50     US-08-723-938-3
51
52     Query Match: 26.8% Score 37, DB 1, Length 395,

```

Best Local Similarity 50.0%; Pred. NO. 80;
Matches 7; Conservative 1; Mismatches 1
07 5 TICTKXXXXVDSG 18
||| |
:| |
DB 272 TLCANGCAALDTG 285

Search completed: May 9, 2003, 16:19:26
Job time : 15.4216 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_gpc4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:15:46 (Search time 15.598 seconds
(without alignments)
218,294 Million cell updates/sec

Title: US-09-851-422a-1
Perfect score: 138
Sequence: 1 GFIATLCITKXXVVDGIDKXXXXXLIQLEKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Minimum Hit 100%

Listing first 45 summaries

Database : Published.Applications.AA:
1: /cgn2_6/p/cdata/2/pubpa/US08_NEM_PUB.pep.*
2: /cgn2_6/p/cdata/2/pubpa/PCT_NEM_PUB.pep.*
3: /cgn2_6/p/cdata/2/pubpa/US06_NEM_PUB.pep.*
4: /cgn2_6/p/cdata/2/pubpa/US07_NEM_PUB.pep.*
5: /cgn2_6/p/cdata/2/pubpa/US07_PUBCOMB.pep.*
6: /cgn2_6/p/cdata/2/pubpa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/p/cdata/2/pubpa/US09_NEM_PUB.pep.*
8: /cgn2_6/p/cdata/2/pubpa/US08_PUBCOMB.pep.*
9: /cgn2_6/p/cdata/2/pubpa/US09_NEM_PUB.pep.*
10: /cgn2_6/p/cdata/2/pubpa/US09_PUBCOMB.pep.*
11: /cgn2_6/p/cdata/2/pubpa/US10_PUBCOMB.pep.*
12: /cgn2_6/p/cdata/2/pubpa/US60_NEM_PUB.pep.*
13: /cgn2_6/p/cdata/2/pubpa/US60_PUBCOMB.pep.*
14: /cgn2_6/p/cdata/2/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score of the result being printed, and is derived by analysis of the score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	31.2	461	US-09-986-480-189	Sequence 189, App
2	43	30.1	460	US-09-972-1124-32	Sequence 6, Appl
3	41	29.7	893	US-09-891-332A-8	Sequence 3, Appl
4	41	29.7	893	US-09-891-332A-7	Sequence 2, Appl
5	41	29.7	893	US-09-891-332A-8	Sequence 8, Appl
6	41	29.7	893	US-09-891-332A-9	Sequence 9, Appl
7	41	29.7	893	US-09-891-332A-9	Sequence 3, Appl
8	41	29.7	893	US-09-891-332A-9	Sequence 1, Appl
9	41	29.7	893	US-09-891-332A-9	Sequence 1, Appl
10	41	29.7	893	US-09-891-332A-9	Sequence 1, Appl
11	40	29.0	718	US-09-928-175-20	Sequence 3, Appl
12	39	28.3	737	US-09-928-175-20	Sequence 2, Appl
13	39	28.3	737	US-09-928-175-20	Sequence 2, Appl
14	39	28.3	737	US-09-928-175-20	Sequence 2, Appl
15	37	26.8	48	US-09-864-761-48379	Sequence 48379, A
16	37	26.8	172	US-09-860-670-112	Sequence 112, Appl
17	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
18	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
19	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
20	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
21	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
22	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
23	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
24	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
25	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
26	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
27	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
28	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
29	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
30	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
31	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
32	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
33	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
34	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
35	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
36	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
37	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
38	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
39	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
40	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
41	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
42	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
43	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
44	37	26.8	172	US-09-860-670-112	Sequence 87, Appl
45	37	26.8	172	US-09-860-670-112	Sequence 87, Appl

20	37	26.8	390	US-09-965-384-15	Sequence 15, Appl
21	37	26.8	395	US-10-094-080-3	Sequence 3, Appl
22	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
23	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
24	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
25	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
26	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
27	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
28	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
29	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
30	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
31	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
32	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
33	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
34	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
35	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
36	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
37	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
38	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
39	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
40	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
41	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
42	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
43	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
44	37	26.8	433	US-09-964-899-11	Sequence 11, Appl
45	37	26.8	433	US-09-964-899-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-09-986-480-189
Sequence 189, Application US/09986480
GENERAL INFORMATION:
APPLICANT: Rosen et al.
FILE OF INVENTION: 143 Human Secreted Proteins
CURRENT FILING DATE: 2001-11-08
PRIORITY FILING DATE: 2000-05-11/US00/12786
PRIOR APPLICATION NUMBER: US 60/34,068
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 456
SEQ ID NO 189
TITLE: 189
SYNOPSIS: 189
ORGANISM: Homo sapiens
US-09-986-480-189
Query Match 31.2%, Score 43; DB 9; Length 431;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;
DB 13 IATLCITKXXVVDGIDKXXXXXLIQLEKXXXX 45
US-09-891-332A-8
Sequence 8, Appl
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 33

```

Matches          9:  Conservative         3:  Mismatches          4:  Indels          1:  Gaps
Oy              17 FGIIDXXXXXLI-OLIID 32
               ||||| | | | | |
Db             287 FGLDKFSQALVMOVLVED 303

RESULT 4
US-09-891-332A-2
Sequence 2, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Inventor(s) : Solus, Joseph
            Yang, Shuwel
TITLE OF INVENTION: Polymersases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C
CITY: New York Ave., N.W., Suite 600
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows NTOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27 Jun-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
PRIORITY DATE: 07-FEB-1997
APPLICATION NUMBER: US 60/037,293
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,4230002
TELEPHONE/FAX/COMMUNICATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOG: none
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-891-332A-2

Query Match
Best Local Similarity 25.8% Pred. No. 55:
Matches      8:  Conservative      6:  Mismatches          15:  Indels          0:  Gaps          0:
Oy           3 INTLCKXXXXXLRDIXKXXXXXLIOLJEDK 33
||||| : : : | : : : : ||||: ||
Db        117 IATLVRAHFLMEFLITJCDKMQLGVNNEK 147

RESULT 5
US-09-891-332A-7
Sequence 7, Application US/09891332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Inventor(s) : Solus, Joseph
            Yang, Shuwel
```

TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof

COUNTRY: USA
ZIP: 20005-9934

COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/89J.332A
FILING DATE: 27-Jun-2001

CLASSIFICATION: <unknown>
PRIORITY DATE: 27-Jun-2001
PRIORITY APPLICATION NUMBER: 09/019,160

FILING DATE: <unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
EXPIRATION DATE: 12-MAR-2002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRUCTURE: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-89J-332A-7

Query Match 29.7% Score 41: DB 9: Length 893:
Best Local Similarity 25.8% Pred. No. 55:
Matches 8; Conservative 6; Mismatches 15: Indels 0: Gaps 0:

QY 3 IATLCYXXHLPDRIKXXXXXLIIDFK 33
||||| : : | : ||||| : :
Db 117 IATLAVRAAFIRLFSLTGDKMDLVNER 147

RESULT 6
US-09-89J-332A-8
US Sequence 8 Application US /0989J332A
Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
SOLUS, Joseph
Yand, Stuart
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: STRINE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-9934
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
COMPUTER READABLE FORM:
MEDIA TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/09/891,332A
 APPLICATION NUMBER: US/09/891,332A
 FILING DATE: 27 JUN 2001
 CLASSIFICATION: <unknown>
 PRIORITY DATE: 09/01/1990
 APPLICATION NUMBER: 09/01/1990
 FILING DATE: <unknown>
 APPLICATION NUMBER: US 60/037,393
 FILING DATE: 07-FEB-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 ADDRESS: 10000 WOODBURN RD
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2600
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 TYPE: amino acid
 LENGTH: 893 amino acids
 SOURCE: synthetic
 ORGANISM: <unknown>
 TOPOLOGY: relevant
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 US-09-891-332A-8
 Query Match 29.7% Score 41; DB 9; Length 893;
 Best Local Similarity 25.8% Pct. No. 55;
 Matches 8; Conservative 6; Mismatches 15; Indels 0; Gaps 0
 Oy 3 IATLCIKKXXXXLDFGDKXXXXLLIEDD 33
 Db 117 IATLVRAAFMRFLITGDKNMLQVNRK 147
 US-09-891-332A-9
 RESULT 7
 Sequence 9, Application US/09891332A
 Patent No. US20020168646A1
 GENERAL INFORMATION:
 APPLICANT: Chatterjee, Deb K.
 Solus, Joseph
 Yano, Shuwei
 TITLE OF INVENTION: Polymers for Analyzing or Typing Polymorphic
 NUMBER OF SEQUENCES: Nucleic Acid Fragments and Uses Thereof
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C
 STREET: 1100 New York Ave., N.W., Suite 600
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 POSTAL CODE: 20004
 TELEPHONE: 202-331-1200
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/891,332A
 APPLICATION DATE: 09/01/1990
 CLASSIFICATION: <unknown>
 PRIORITY DATE: 07-FEB-1997
 FILING DATE: 07-FEB-1997
 APPLICATION NUMBER: US 60/037,393
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 ADDRESS: 10000 WOODBURN RD
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2600

```

TELFAK: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
    LENGTH: 893 amino acids
    STRANDNESS: not relevant
    TOPOLOGY: not relevant
    MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-891-332a-9

Query Match
Best Local Similarity 29.7%: Score 41: DB 9: Length 893:
Best Local Similarity 25.8%: Pred. No. 55:
Matches 8: Conservative 15: Indels 0: Gaps 0:

Oy 3 IATCTKXXXXXVLDGIDKXXXXXLTQLLEPK 33
DB 117 IATLVAAARFLARSLITGCRDMQLVNER 147

RESULT 8
US-09-229-173-3
Sequence 3, Application US/09229173
Publication NO. US20030027296A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA Polymers from Thermotoga
TITLE OF INVENTION: Cloned DNA Polymers from Thermotoga
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,173
FILING DATE: 13-JAN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/706,702
FILING DATE: 08-SEP-1996
REGISTRATION NUMBER: 08/689,807
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
    LENGTH: 893 amino acids

```

```

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-229-173-3
Query Match
Best Local Similarity 29.7%: Score 41: DB 9: Length 893:
Best Local Similarity 25.8%: Pred. No. 55:
Matches 8: Conservative 15: Indels 0: Gaps 0:

Oy 3 IATCTKXXXXXVLDGIDKXXXXXLTQLLEPK 33
DB 117 IATLVAAARFLARSLITGCRDMQLVNER 147

RESULT 9
US-09-741-664-1
Sequence 1, Application US/09741664
Patent No. US20010041334A1
GENERAL INFORMATION:
APPLICANT: Smith, Joseph
TITLE OF INVENTION: Stable Compositions for Nucleic Acid
TITLE OF INVENTION: Sequencing and Amplification
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P. L. L. C.
STREET: 1100 New York Avenue, N.W., Suite 600
STATE: Washington
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/741,664
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,021
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
TELEPHONE/DOCKET NUMBER: 0942,3910000
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
    LENGTH: 893 amino acids
    TYPE: amino acid
    STRANDNESS: single
    TOPOLOGY: linear
US-09-741-664-1
Query Match
Best Local Similarity 29.7%: Score 41: DB 10: Length 893:
Best Local Similarity 25.8%: Pred. No. 55:
Matches 8: Conservative 15: Indels 0: Gaps 0:

Oy 3 IATCTKXXXXXVLDGIDKXXXXXLTQLLEPK 33
DB 117 IATLVAAARFLARSLITGCRDMQLVNER 147

RESULT 10
US-09-741-664-2
Sequence 2, Application US/09741664
Patent No. US20010041334A1
GENERAL INFORMATION:

```


APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Zhang, Mengsheng
TITLE OF INVENTION: GENOME-DEIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
TITLE REFERENCE: Aomica-X-1
CURRENT APPLICATION NUMBER: US/09/0654,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: CR 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/226,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661

Search completed: May 9, 2003, 16:20:16
Job-time : 17.598 secs

A:Cross-references: EMBL:X76904; NID:9509761; PIDN:CA54225.1; PID:9509762
 A:Accession: S61453
 A:Residues: 20-63 <LEI>
 C:Superfamily: amoebapore; saposin repeat homology
 C:Keyworts: antibacterial
 F:1-19/Domain: signal sequence #status predicted <SIC>
 F:16-96/Domain: saposin repeat homology <SAP>
 F:20-96/Product: amoebapore B status predicted <NAP>
 F:24-96:27-90:54-63/Distulfide bonds: #status predicted
 Query Match 45.7%: Score 63; DB 1; Length 96;
 Best Local Similarity 43.8%: Pred. No. 0.00049;
 Matches 14; Conservative 7; Mismatches 3; Indels 2;

DB 1 GFATLCTKXXXXVLDGIDXXXXXLIOLIED 32
 59 GFLTLCTCK-----ILSTFVDE---LVALLIEN 82

RESULT 3

amoebapore C precursor - Entamoeba histolytica
 C:Species: Entamoeba histolytica
 C:Date: 16-Feb-1993 #sequence_revision 12-May-1995 #texL_change 18-Jun-1999
 F:1-24/Domain: signal sequence #status predicted <SIC>
 R:Leidner, M.; Andre, J.; Nickel, R.; Tamlich, E.; Mueller-Berhard, H.J.
 MOL. Microbiol. 14, 895-904, 1994
 A:Title: Amoebapores, a family of membranolytic peptides from cytoplasmic granules of *Et*
 A:Reference number: S61439; MUID:95231296; PMID:7715451
 A:Accession: S61439
 A:Molecule type: DNA
 A:Cross-references: EMBL:X76903; NID:9509763; PIDN:CA54225.1; PID:9509764
 A:Accession: S61453
 A:Residues: 25-69 <LEI>
 A:Molecule type: protein
 C:Superfamily: amoebapore; saposin repeat homology
 C:Keyworts: antibacterial
 F:1-24/Domain: signal sequence #status predicted <SIC>
 F:25-101/Product: amoebapore C status predicted <NAP>
 F:25-101/Product: amoebapore C status predicted <NAP>
 F:29-101:32-95:55-70/Distulfide bonds: #status predicted

Query Match 39.1%: Score 54; DB 1; Length 101;
 Best Local Similarity 35.7%: Pred. No. 0.022;
 Matches 10; Conservative 7; Mismatches 7; Indels 4; Gaps 1;
 DB 1 GFATLCTKXXXXVLDGIDXXXXXLIOLIED 32
 64 GVLCTCK-----IYSIGDILKIEKIE 87

RESULT 4

hypothetical protein C10618 [imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #texL_change 03-Jun-2002
 C:Accession: AB1410
 R:Perhalla, J.; Wren, B.W.; Mungall, K.; Kelly, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W.; Quail, M.; Mungall, K.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrett, N.; et al.
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hyf
 A:Accession: AB1410
 A:Residues: 1-1215 <P&R>
 A:Cross-references: GB:AL119075; GB:AL111168; NID:96967817; PIDN:CA875254.1; PID:9696808
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics: C10618

Query Match 31.9%: Score 44; DB 2; Length 215;
 Best Local Similarity 40.0%: Pred. No. 3.2;
 Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 DB 1 GFATLCTKXXXXVLDGID 20
 8 CFLANCKKIAWPSIDPIID 27

RESULT 5

hemagglutinin - phocine distemper virus
 C:Species: Phocine distemper virus
 C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #texL_change 16-Jul-1999
 C:Accession: J01369
 R:Koenigsmann, J.; Blixenkrone-Møller, M.; Sharma, B.; Oerwell, C.; Norrby, E.
 J. Gen. Virol. 72, 2959-2966, 1991
 A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglut
 A:Reference number: J01369; MUID:92113538; PMID:1765768
 A:Accession: J01369
 A:Molecule type: genomic RNA
 A:Residues: 1-607 <KOV>
 A:Cross-references: GB:X26979; NID:9536790; PIDN:CA85428.1; PID:9536791
 C:Genetics: C10618

C:Superfamily: measles virus hemagglutinin
 C:Keyworts: glycoprotein; glycoprotein; glycoprotein; glycoprotein
 F:39-55/Domain: transmembrane status predicted <TM>
 F:19:149:276:391:422:456:587/Binding site: carbohydrate (Asn) (covalent) #status pred
 Query Match 31.9%: Score 44; DB 1; Length 607;
 Best Local Similarity 28.6%: Pred. No. 9.4;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 DB 292 LASLCTKRESTITLLNCPRESONSVLAVI 319

RESULT 6

hemagglutinin - phocine distemper virus (strain Uster/88)
 C:Species: Phocine distemper virus
 N:Alternative names: attachment protein
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #texL_change 09-Sep-1994
 C:Accession: J01535
 R:Curran, M.D.; O'Loan, D.; Kennedy, S.; Rima, B.K.
 J. Gen. Virol. 73, 1189-1194, 1992
 A:Title: Molecular characterization of phocine distemper virus: Gene order and sequen
 A:Accession: J01535; MUID:92268877; PMID:1586532
 A:Residues: 1-607 <CUR>
 A:Molecule type: mRNA
 A:Cross-references: GB:D10371
 C:Genetics: C10618

C:Superfamily: measles virus hemagglutinin
 C:Keyworts: glycoprotein; glycoprotein; glycoprotein; glycoprotein
 F:39:55/Domain: transmembrane status predicted <TM>
 F:19:149:276:391:422:456:587/Binding site: carbohydrate (Asn) (covalent) #status pred
 Query Match 31.9%: Score 44; DB 1; Length 607;
 Best Local Similarity 28.6%: Pred. No. 9.4;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 DB 292 LASLCTKRESTITLLNCPRESONSVLAVI 319

RESULT 7

hypothetical protein C44B1.3 - *Caenorhabditis elegans*

Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 OY 18 CIDKXXXXXLIQLEDK 33
 DB 363 GIDKKKIVITELPEEK 378

RESULT 12

serine protease MucD precursor PA0766 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #extl_change 03-Jun-2002
 C:Species: Pseudomonas aeruginosa
 C:Accession: F83550
 R:Cover: C.K.; Pharm.; X.Q.; Ervin, A.B.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yum, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,
 .; Lohy, S.; Olson, M.V.
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 gen.
 A:Reference number: A82950; MID:20437337; PMID:10984043
 A:Accession: F83550
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-474 <S>
 A:Cross-references: CB:AE004511; CB:AE004091; NID:95946646; PIDN:AA04135.1; GSPDB:GN001
 A:Experimental source: strain PAO1
 C:Genetic:
 A:Gene: mucD; PA0766
 C:Superfamily: Helicobacter serine protease

Query Match 30.1%; Score 41.5; DB 2; Length 474;
 Best Local Similarity 32.9%; Pred. No. 21;
 Matches 3; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
 OY 17 FGIDKXXXXLI-QLIED 32
 DB 287 FGIDKPSGALVALVED 303

RESULT 13

hypothetical protein PYF [imported] - Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #extl_change 15-Jun-2001
 C:Accession: H90208
 R:Shen, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Chan-
 Jong, I.; Jettles, A.C.; Kozera, C.W.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R
 .; submitted to Genbank May 11, 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: H90208
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-222 <S>
 A:Cross-references: CB:AE006641; NID:913813783; PIDN:AAK40927.1; GSPDB:GN00155
 C:Genetic:
 A:Gene: PYF
 C:Superfamily: hypothetical protein M0252

Query Match 29.7%; Score 41; DB 2; Length 222;
 Best Local Similarity 36.2%; Pred. No. 12;
 Matches 3; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 OY 14 VIDGIDKXXXXXLIQL 29

DB 36 VIDLSVYKTRRLIGL 51

RESULT 14

OIL protein - vaccinia virus (strain Copenhagen)
 C:Species: vaccinia virus

A:Note: host Homo sapiens (man)
 C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #extl_change 20-Jun-2000
 C:Accession: D42510
 R:Submitted to Genbank, June 1990
 A:Reference number: A33172
 A:Accession: D42510
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-666 <S>
 C:Superfamily: variola major virus hypothetical protein Q1L

Query Match 29.7%; Score 41; DB 2; Length 666;
 Best Local Similarity 31.0%; Pred. No. 36;
 Matches 9; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
 OY 4 ATLCKXXXXXVLPDGIKXXXXXLIQLEID 32

DB 129 ATYLKINMYMINKIDVYDEIDLYMD 157

RESULT 15

hypothetical protein Q1L - variola major virus
 C:Species: variola major virus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #extl_change 21-Jul-2000
 C:Accession: F78491
 R:Massung, R.F.; Eposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Au
 Nature 366, 748-751, 1993
 A:Title: Potential virulence determinants in terminal regions of variola smallpox vir
 A:Reference number: Z20488; MID:94088747; PMID:8264796
 A:Accession: F78491
 A:Status: preliminary; translated from CB/EMBL/DD83
 A:Residues: 1-666 <S>
 A:Cross-references: EMBL:U22579; NID:9623595; PIDN:AA60801.1; PID:9438971
 A:Experimental source: strain Bangladesh-1975
 C:Superfamily: variola major virus hypothetical protein Q1L

Query Match 29.7%; Score 41; DB 2; Length 666;
 Best Local Similarity 31.0%; Pred. No. 36;
 Matches 9; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
 OY 4 ATLCKXXXXXVLPDGIKXXXXXLIQLEID 32
 DB 129 ATYLKINMYMINKIDVYDEIDLYMD 157

Search completed: May 9, 2003, 16:18:44
 Job time: 26.2157 secs

GenCore version 5.1.4-p3.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:08:20 / Search time 11.9706 seconds

(without alignments) 128.200 Million cell updates/sec

Title: US-09-851-422a-1

Sequence: 1 GFATLCTKXXVDFGIDKXXXXLQLLEKXXXX 37

Scored table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SWSsprotL40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being predicted, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	102	73.9	98	1	PPPA-ENTH1
2	101	73.2	97	1	PPPA-ENTH1
3	101	73.2	97	1	PPPA-ENTH1
4	54	39.1	101	1	PPPC-ENTH1
5	43	31.2	594	1	STBL-HUMAN
6	43	31.2	594	1	STBL-MOUSE
7	41.5	30.1	915	1	SVR-METKA
8	41	29.7	222	1	PYRF-SULSO
9	41	28.7	666	1	VOOL-VANCC
10	41	28.7	666	1	VOOL-VANCC
11	40	28.6	1612	1	TP2B-MOUSE
12	40	28.6	1612	1	TP2B-MOUSE
13	39.5	28.3	185	1	VHR2-SHIVA
14	39	28.3	185	1	VHR2-SHIVA
15	39	28.3	290	1	VG33-HSV11
16	39	28.3	341	1	PLSX-YIEBH
17	39	28.3	341	1	PLSX-YIEBH
18	39	28.3	341	1	PLSX-YIEBH
19	39	28.3	341	1	PLSX-YIEBH
20	39	28.3	341	1	PLSX-YIEBH
21	39	28.3	341	1	PLSX-YIEBH
22	39	28.3	341	1	PLSX-YIEBH
23	39	28.3	341	1	PLSX-YIEBH
24	39	28.3	341	1	PLSX-YIEBH
25	39	28.3	341	1	PLSX-YIEBH
26	39	28.3	341	1	PLSX-YIEBH
27	39	28.3	341	1	PLSX-YIEBH
28	39	28.3	341	1	PLSX-YIEBH
29	39	28.3	341	1	PLSX-YIEBH
30	39	28.3	341	1	PLSX-YIEBH
31	39	28.3	341	1	PLSX-YIEBH
32	39	28.3	341	1	PLSX-YIEBH
33	39	28.3	341	1	PLSX-YIEBH

RESULT 1
PPPA-ENTH1 STANDARD: PRT; 98 AA.
ID: P34095 Entamoeba h
101-PPB-1994 (Ref. 28, Created)
01-FEB-1994 (Ref. 28, Last sequence update)
16-OCT-2001 (Ref. 40, Last annotation update)
DE Pore-forming peptide amebapore A precursor (EH-APP).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
NCBI_TaxID=5759;
SEQUENCE FROM N.A., AND SEQUENCE OF 22-46.
RC STRAIN-HM-1:IMSS;
MEDLINE=94099892; PubMed=8274224;
Stratus I., Ielppe M., Loutens C., Tannich E.;
Bruscia R.E., Tannich E., Nickel R., van der Goot G., Patus F.,
Horstmann R.D., Mueller-Eberhard H.J.;
Primary and secondary structure of the pore-forming peptide of
the protozoan parasite Entamoeba histolytica.
PMID=94099892.
RT EMBL J. 11:35013506(1992).
RM
PP
SEQUENCE FROM N.A.
RC STRAIN-HM-1:IMSS;
MEDLINE=94099892; PubMed=8274224;
Stratus I., Ielppe M., Loutens C., Tannich E.;
Bruscia R.E., Tannich E., Nickel R., van der Goot G., Patus F.,
Horstmann R.D., Mueller-Eberhard H.J.;
Primary and secondary structure of the pore-forming peptide of
the protozoan parasite Entamoeba histolytica.
PMID=94099892.
RT EMBL J. 11:35013506(1992).
RM
PP
PARTIAL SEQUENCE.
RC STRAIN-HM-1:IMSS;
MEDLINE=95231296; PubMed=7715431;
Lelpe M., Andrej J., Nickel R., Tannich E., Mueller-Eberhard H.J.;
Amebapores, a family of membranolytic peptides from cytoplasmic
granules of Entamoeba histolytica: Isolation, primary structure, and
Mol. Microbiol. 14:1895-904(1994).
-1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTENS, NO ACTIVITY AGAINST
E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
-1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
-1- SIMILARITY: CONTAINS 1 SAPOTIN B-TYPE DOMAIN.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by for commercial
entities requires a license agreement (See <http://www.isb.slb.ch/announce/>)

RA MEDLINE-95231296; PubMed-7715451;
 RA Lelepe M., Andre J., Nickel R., Taniich E., Mueller-Eberhard H.J.:
 RA Aminoacyl-tRNA synthetase: a family of membrane proteins from cytoplasmic
 RA and mitochondrial membranes. The aminoacyl-tRNA synthetase family structure, and
 RA pore formation in bacterial cytoplasmic membranes.*;
 RL Mol. Microbiol. 14:895-904(1994).
 CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
 CC HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS. NO ACTIVITY AGAINST
 CC E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
 CC -1- SIMILARITY: CONTAINS 1 SAPODIN B-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X76904; GMS43225.1;
 DR InterPro: IPR000004; Sapp.
 DR SMART: SM00118; Sapp; 1.
 KW Signal; Anticidetic.
 FT SIGNAL 24
 FT PEPTIDE 25 101 PORE-FORMING PEPTIDE AMEOPAPORE C.
 FT DISULFID 32 95 BY SIMILARITY.
 FT DISULFID 59 70 BY SIMILARITY.
 SO SEQUENCE 101 AA; 10855 MW; 28AC72135E73439B CRC64;
 Query Match 39.18; Score 54; DB 1; Length 101;
 Best local Similarity 35.78; Pred. No. 0.0064;
 Matches 10; Conservative 7; Mismatches 7; Indels 4; Gaps 1;
 Oy 1:|||||
 Db 64 GIVETCTK---IVSYGIDLEIKLE 87
 RESULT 5
 STR1_HUMAN STANDARD; PFM; 554 AA.
 AC 064130; 062759; 028208;
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 01-NOV-1997 (Rel. 35; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Syntaxin binding protein 1 (Unc-18 homolog) (Unc-18A) (Unc-18-1) (N-
 DE Sec1) (resc1) (p87).
 DN SYNTAXIN OR UNC18A.
 OS Rattus norvegicus (Rat), and
 OS Rattus norvegicus (Rat), and
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NX NCBI_TaxID:9606; 10116; 9913;
 RX MEDLINE-95231296; PubMed-7715451;
 RX Lelepe M., Andre J., Nickel R., Taniich E., Mueller-Eberhard H.J.:
 RX Aminoacyl-tRNA synthetase: a family of membrane proteins from cytoplasmic
 RX and mitochondrial membranes. The aminoacyl-tRNA synthetase family structure, and
 RX pore formation in bacterial cytoplasmic membranes.*;
 RL Mol. Microbiol. 14:895-904(1994).
 CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
 CC HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS. NO ACTIVITY AGAINST
 CC E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
 CC -1- SIMILARITY: CONTAINS 1 SAPODIN B-TYPE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X76904; GMS43225.1;
 DR InterPro: IPR000004; Sapp.
 DR SMART: SM00118; Sapp; 1.
 KW Signal; Anticidetic.
 FT SIGNAL 24
 FT PEPTIDE 25 101 PORE-FORMING PEPTIDE AMEOPAPORE C.
 FT DISULFID 32 95 BY SIMILARITY.
 FT DISULFID 59 70 BY SIMILARITY.
 SO SEQUENCE 101 AA; 10855 MW; 28AC72135E73439B CRC64;

RA [3]
 RA SEQUENCE FROM N.A.
 RA SPECIES-Human; TISSUE-Skin;
 RA Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY PARTICIPATE IN THE REGULATION OF SYNAPTIC
 CC DOCKING AND FUSION, POSSIBLY THROUGH INTERACTION WITH GTP-BINDING
 CC PROTEINS. ESSENTIAL FOR NEUROTRANSMISSION AND BINDS SYNTAXIN, A
 CC COMPONENT OF THE SYNAPTIC VESICLE FUSION MACHINERY PROBABLY IN A
 CC 1:1 RATIO. CAN INTERACT WITH SYNTAXIN 1, 2, AND 3 BUT NOT
 CC SYNTAXIN 4. MAY PLAY A ROLE IN DETERMINING THE SPECIFICITY OF
 CC SYNAPTIC VESICLE FUSION.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, 1/4 (shown here) and 2/4B; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN AND SPINAL CORD. HIGHLY ENRICHED IN
 CC AXONS.
 CC -1- DEVELOPMENTAL STAGE: PAINT LEVELS ARE DETECTABLE AT EMBRYONIC DAY
 CC 10.5. PAINT LEVELS AT LATER EMBRYONIC AGES AND PEAKING AT
 CC POSTNATAL DAY 7.
 CC -1- SIMILARITY: BELONGS TO THE STRXP/UNC-18/SEC1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 RA [3]
 RA SEQUENCE FROM N.A.
 RA SPECIES-Human; TISSUE-Skin;
 RA Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY PARTICIPATE IN THE REGULATION OF SYNAPTIC
 CC DOCKING AND FUSION, POSSIBLY THROUGH INTERACTION WITH GTP-BINDING
 CC PROTEINS. ESSENTIAL FOR NEUROTRANSMISSION AND BINDS SYNTAXIN, A
 CC COMPONENT OF THE SYNAPTIC VESICLE FUSION MACHINERY PROBABLY IN A
 CC 1:1 RATIO. CAN INTERACT WITH SYNTAXIN 1, 2, AND 3 BUT NOT
 CC SYNTAXIN 4. MAY PLAY A ROLE IN DETERMINING THE SPECIFICITY OF
 CC SYNAPTIC VESICLE FUSION.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms, 1/4 (shown here) and 2/4B; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN AND SPINAL CORD. HIGHLY ENRICHED IN
 CC AXONS.
 CC -1- DEVELOPMENTAL STAGE: PAINT LEVELS ARE DETECTABLE AT EMBRYONIC DAY
 CC 10.5. PAINT LEVELS AT LATER EMBRYONIC AGES AND PEAKING AT
 CC POSTNATAL DAY 7.
 CC -1- SIMILARITY: BELONGS TO THE STRXP/UNC-18/SEC1 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to: license@isb-sib.ch).

CC BML: D63851: BAA19483.1: -
 CC BML: AF004562: AAC39688.1: -
 CC BML: BC015749: AAB15749.1: -
 CC BML: L26087: AAB1112.1: -
 CC BML: L26087: AAB1112.1: -
 CC BML: U28248: AAB19246.1: -
 CC BML: U28248: AAB19246.1: -
 CC BML: U21116: AAB6350.1: -
 CC BML: AF153327: AAD37018.1: -
 CC Genew: HGNC:114444: STXB1.
 DR MIM: 602926: -
 DR InterPro: IP001619: Sec1-1like.
 DR Pfam: PF00995: Sec1.1.
 DR Protein Transport: Alternative splicing.
 DR VARSIB: 576 594
 FT SPROCK: 0.58 (IN PROGRESS)
 FT CONFLICT 291 291 L > V (IN REF. 4: ABA1112).
 FT SEQUENCE 594 AA: 67568 MW: 200725875CE0F3 CRC64:
 Query Match 31.2% Score 43: DB 1: Length 594:
 Best Local Similarity 45.5%: Pred. No. 4.2:
 Matches 19: Conservative 1: Mismatches 15: Indels 2: Gaps 1:
 Oy 3 INACT--KXXXVLDGIDKXXXXXLIQLEK 33
 DB 176 IATLCATLKEPAVHYRGEDNALALQLODK 208
 ID STRL.MOUSE STANDARD: PRT: 594 AA.
 AC 008599:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DT SYNTAXIN binding protein 1 (unc-18 homolog) (unc-18-1).
 OR Mus musculus (mouse)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCL:TaxID-10090:
 RN [1]
 RX SEQUENCE FROM N.A. SRA:
 RX SRR15642.67.138683.83.10.
 RA Genovese-Ando K., Kishiyama H., Wakiada M., Ikawa Y.:
 "A murine neural-specific homolog corrects cholinergic defects in *Cnch* mutants." *Neuron* 16:695-6702(1996).
 RL J. Neurosci. 16:695-6702(1996).
 CC FUNCTION: MAY PARTICIPATE IN THE REGULATION OF SYNAPTIC VESICLE DOCKING AND FUSION, POSSIBLY THROUGH INTERACTION WITH GTP-BINDING COMPONENT OF THE SYNAPTIC VESICLE FUSION MACHINERY. A
 CC 1.1 RATIO. CAN INTERACT WITH SYNTAXIN 1, 2, AND 3 BUT NOT
 CC SYNTAXIN 4. MAY PLAY A ROLE IN DETERMINING THE SPECIFICITY OF INTRACELLULAR FUSION REACTIONS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SYMB/DNC-18/SEC1 FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL, distribution is by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to: license@isb-sib.ch).

CC BML: D45031: BAA19479.1: -
 CC BML: D45031: BAA19479.1: -
 DR InterPro: IP001619: Sec1-like.

DR Pfam: PF00995: Sec1.1.
 DR Protein Transport: Protein transport.
 DR SEQUENCE 594 AA: 67701 MW: 6367628215B3E847 CRC64:
 Query Match 31.2% Score 43: DB 1: Length 594:
 Best Local Similarity 45.5%: Pred. No. 4.2:
 Matches 19: Conservative 1: Mismatches 15: Indels 2: Gaps 1:
 Oy 3 INACT--KXXXVLDGIDKXXXXXLIQLEK 33
 DB 176 IATLCATLKEPAVHYRGEDNALALQLODK 208
 ID SYA.METKA STANDARD: PRT: 915 AA.
 AC 087471:
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA Ligase) (Alans).
 OS ALAS OR MK0900.
 CC Methanopyrus kandleri.
 CC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 CC Methanopyrus.
 CC NCL:TaxID-2320:
 RX SEQUENCE FROM N.A. SRA:
 RX STRAIN-AV19 / DSM 6324 / JCM 9639:
 RX MEDLINE-21927647: PubMed-11530014:
 RA Shestarev A.I., Mezheva K.V., Makrova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Matusev R.L., Wolf Y.I., Stettler K.O.,
 RA "Phylogenetic analysis of the genome of the hyperthermophilic
 .RT and monophyly of archaeal methanogens".
 CC Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 CC -1- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
 CC di-phosphate + L-alanyl-tRNA(Ala).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL, distribution is by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to: license@isb-sib.ch).

CC BML: AE010379: AAM02113.1: ALT INT.
 CC DR PROSITE: PS50860: AL-TRNA-LIGASE-II.ALA: 1.
 RM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 RM Complete proteome.
 SQ SEQUENCE 915 AA: 103508 MW: 60D8BA1935C44E CRC64:
 Query Match 30.1% Score 41.5: DB 1: Length 915:
 Best Local Similarity 29.0%: Pred. No. 12:
 Matches 9: Conservative 9: Mismatches 12: Indels 1: Gaps 1:
 Oy 1 GPTATCTKXXXVLDGIDKXXXXXLIQLE 31
 DB 355 GTLARLAVIRALRLD-GLDREYLLVEE 384
 ID PYRF.SULSO STANDARD: PRT: 222 AA.
 AC 090X10:
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Oxidolide 5'-phosphatase decarboxylase (EC 4.1.1.23) (OMP decarboxylase)

FT	TRANSMEM	359	376	POTENTIAL.
FT	DOMAIN	377	920	CYTOLASMIC. (POTENTIAL).
FT	TRANSMEM	621	920	POTENTIAL.
FT	TRANSMEM	921	940	POTENTIAL.
FT	TRANSMEM	955	974	CYTOLASMIC. (POTENTIAL).
FT	DOMAIN	975	1004	CYTOLASMIC. (POTENTIAL).
FT	TRANSMEM	1005	1027	POTENTIAL.
FT	DOMAIN	1028	1040	EXTRACELLULAR. (POTENTIAL).
FT	TRANSMEM	1041	1063	POTENTIAL.
FT	DOMAIN	1064	1069	CYTOLASMIC. (POTENTIAL).
FT	TRANSMEM	1070	1090	POTENTIAL.
FT	TRANSMEM	1091	1100	CYTOLASMIC. (POTENTIAL).
FT	TRANSMEM	1108	1132	POTENTIAL.
FT	DOMAIN	1133	1189	CYTOLASMIC. (POTENTIAL).
FT	MOD. RES	424	434	PHOSPHORYLATION. (BY SIMILARITY).
FT	METAL	865	885	MANGNESEUM (BY SIMILARITY).
FT	METAL	869	869	MANGNESEUM (BY SIMILARITY).
FT	SEQUENCE	1189	135309	MM: DCP95ICFE353583 CR664;

```

Oy      5 TUC7XXXXXXLPSIDKXXXXLITD---EDK 33
Db      670 TLTETDRHLLIMAAKKEKDLHLSSTAVADK 702

RESULT 12
TP2B_MOUSE STANDARD: PRT; 1612 AA.
AC      064511:
DT      15-JUL-1998 (rel. 36, Created)
DT      15-JUL-1998 (rel. 36, Last annotation update)
DT      15-JUN-2002 (rel. 41, Last annotation update)
DI      DAA Topoisomerase II, beta Isozyme (EC 5.99.1.3).
DP      top2b
OS      Mus musculus (Mouse)
OC      Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX      NCBI_TaxID=10090:
RN      11)
RP      SEQUENCE FROM N.A.
RA      STRAIN=BALE/c; TISSUE=Brain;
RA      Migrate.N.; Addl.N.A.: K18001.A.;
RA      Migrate.N.; Addl.N.A.: Germany/ODM databases
RI      FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
RI      BREAKS AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
RI      MAKES DOUBLE-STRAND BREAKS.
-1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
-1- of double-stranded DNA.
-1- SUBUNIT: HOMODIMER (BY SIMILARITY).
-1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
-1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
-1- NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
-1- EXCLUSIVELY RELAX POSITIVE SUPERCOILS. II TOPOISOMERASE FAMILY.
-1- SIMILARITY: 100% IDENTITY WITH TOPOISOMERASE II.
CC      This SNIS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed, usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce)
CC      or send an email to license@sib-sib.ch).
CC
CC      P804: P8047226.1;
CC      HSP: P05746; IBCW
DR      MGI: 98791; Top2b.
DR      InterPro: IPRO03594; ATPbind_ATPase.
DR      InterPro: IPRO03597; CBFA_NTF2_LoPbL.
DR      InterPro: IPRO01245; DNA_TopoIsmII.
DR      InterPro: IPRO02205; DNA_TopoIsmII.

```


[illegible]

```
CC --
DR EMBL: AF06871. AAC03106.1; -
DR HSSP: P24293.1HR.
DR PROSITE: PS01371. RUVG.1.
DR PRINTS: PR00696. RSQVYASEMUUC.
DR PRODOM: PD008333. RUVG.1.
DR TIGRFSMS: TIGR00228. RUVG.1.
DR PROSITE: PS01371. RUVG.1.
KW Hydrolase; Nuclease; DNA repair; DNA recombination; Magnesium; Complete proteome.
FT METAL 76 76 MAGNESIUM (BY SIMILARITY).
FT METAL 148 148 MAGNESIUM (BY SIMILARITY).
FT METAL 151 151 MAGNESIUM (BY SIMILARITY).
FT FT 151 151 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 163 AA: 1740 MW: 5024A631B76CDD CRC64:
Query Match Best Local Similarity 33.3%; Pred. No. 5.1; Length 163;
Matches 11; Conservative 2; Mismatches 7; Gaps 1.
Dy 1 GFATLCF-----KXXXYDFGIKKXXYY 26
Db 16 GFGATCTPELVOSTYKNVNDYGVFKTSADI 48
WVRSVF 14
ID VNR2.SPVVA STANDARD: PRT: 185 AA.
CD VNR2.SPVVA
AC P23333.
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Probable host range protein 2.
DS SWEBA.
OS Swinepox virus (strain Kazda) (SPV)
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae; Suipoxvirus.
NC NC_017412.1
NM NC_017412.1
RP MEDLINE:91195285; PubMed:1840707;
RA SCHUTTELMAN W.M., FRITZLEY D.N.;
RL of swinepox virus".
RL Virology 181:727-732(1991).
CC -!- FUNCTION: Plays a role for multiplication.
CC -!- different cell types (by similarity).
CC -!- SIMILARITY: BELONGS TO THE POXVIRUSES C7 FAMILY.
CC CC This SWISS-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL consortium. The European Bioinformatics Institute has made this release as its content is in no way derived from EMBL. The EMBL accession number is X15111. This release is in no way derived from EMBL. The EMBL accession number is X15111. This release is in no way derived from EMBL. The EMBL accession number is X15111.
CC CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/cc or send an email to licenses@isb-sib.ch).
CC EMBL: M55931. AAA41892.1; -.
DR EMBL: M55931. AAA41892.1; -.
DR InterPro: IPR004507; Pox_C7_PBA.
DR Pfam: PF03287; Pox_C7_PBA.1.
DR Early protein.
SQ SEQUENCE 185 AA: 21751 MW: 46D65F31DB63C56 CRC64:
Query Match Best Local Similarity 28.3%; Score 39; DB 1; Length 185;
Matches 7; Conservative 5; Mismatches 10; Indels 0; Gaps 0.
Dy 7 CRKXXVYLDFGICRXXXLLIO 28
Db 106 CRKEKPINIMINKKXEIE 127
```

```

RESULT 15
V33_LSV11
ID KC33_HSV11 STANDARD: PRT: 290 AA.
AC 000118:
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Hypothetical gene 33 protein.
OR 33 kDa putative herpesvirus 1 (Chamoni catfish virus) (CCV).
OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RM 111
RP SEQUENCE FROM N.A.
RC STRAIN=Aburn 1;
RC MEDLINE=92087490; PubMed=1727613;
RK 111
RP "Chamoni catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, but the EMBL/NCBI disclaimer and the copyright notice must be included
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL, M75136; AAA88136.1; .
DB FRC, G36789; G36789.
DB HYPHERS13; HYPHERS13.
SQ SEQUENCE 290 AA. 34173 MW: 21C26260985A7E5E CXC64;
Query Match 28.3%; Score 39; DB 1; Length 290;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GFIATCTK 9
DB 217 GFIITCTK 225

```

Search completed: May 9, 2003, 16:16:16
 Job time : 14.9706 secs

Q9UBG4
ID Q9UBG4 PRELIMINARY: PRT: 101 AA.
AC 09UBG4
DC 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pore-forming protein isoform C precursor.
GN Pp-C.
OS Entamoeba dispar.
OC Eukaryota; Eumetazoa; Entamoeba.
NC [1]_TextID=46681;
RN [1]_TextID=46681;
RP SEQUENCE FROM N.A.
RC STRAIN=SAW 142;
RX MEDLINE=99443604; PubMed=10518795;
RA Nickel R., Olt C., Dandekar T., Laipke M.;
AB "Pore-forming peptides of entamoeba dispar similarity and divergence
in amoebozoa: In structure, expression and activity."
PMBL: A9802528; A9804196.1.
DR InterPro: IPR000004; SAPB.
DR Prodom: PD001732; SAPB: 1.
DR SMART: SM00118; SAPB: 1.
KT SIGNAL.
FT SIGNAL. 1 24
LN LINEAR POTENTIAL.
NM NUCLEONIC ACID SEQUENCE FROM N.A.
NC [1]_TextID=46681;
RN [1]_TextID=46681;
RP SEQUENCE 101 AA; 10877 MW; 7699A7E5D8078B CMC64;
SC

Query Match 35.5%; Score 49; DB 5; Length 101;
Best Local Similarity 35.7%; Pred. No. 0.22;
Matches 10: Conservative 7; Mismatches 7; Indels 4; Gaps 1;
Db 64 CIVETVCSK---IVSYGDKLLEKITE 87

RESULT 3
ID 015591 PRELIMINARY: PRT: 87 AA.
AC 015591
DC 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Amoebozoa C homologue (Fragment).
OS Entamoeba histolytica.
OC Eukaryota; Eumetazoa; Entamoeba.
NC [1]_TextID=5759;
RN [1]_TextID=5759;
RP SEQUENCE FROM N.A.
RC STRAIN=HM-1;
RX MEDLINE=9736144; PubMed=9245698;
RA Tanaka T., Tanaka M., Mitsui Y.;
AB "Analysis of expressed sequence tags (ESTs) of the parasitic protozoa
Entamoeba histolytica." Commun. 236:611-615(1997).
PMBL: A9802528; A9804196.1.
DR InterPro: IPR000004; SAPB.
DR SMART: SM00118; SAPB: 1.
KT NON_TER 1 1
LN LINEAR POTENTIAL.
NM NUCLEONIC ACID SEQUENCE FROM N.A.
NC [1]_TextID=5759;
RN [1]_TextID=5759;
RP SEQUENCE 87 AA; 9131 MW; FDDIC1H0FA8B06E CMC64;
SC

Query Match 32.3%; Score 46; DB 5; Length 87;
Best Local Similarity 32.4%; Pred. No. 0.71;
Matches 10: Conservative 4; Mismatches 9; Indels 4; Gaps 1;
Db 59 GLAVTLCTP---IASFCIAKLASTL 81

Q9PH93
ID Q9PH93 PRELIMINARY: PRT: 215 AA.
AC 09PH93
DC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Hypothetical protein Cj0618.
GN Cj0618.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group.
NC [1]_TextID=197;
RN [1]_TextID=197;
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Mren B.W., Mungall K., Kelsey J.M., Churcher C.,
BA Basham D., Chillingworth T., Davies R.H., Delwail P., Dooly S.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrett B.C.;
AB "The genome sequence of the food-borne pathogen Campylobacter jejuni
reveals hypervariable sequences." Nature 403:665-668(2000).
PMBL: A1130075; C987523.1.
DR EMBL: AL130075; C987523.1.
NC [1]_TextID=197;
RN [1]_TextID=197;
RP SEQUENCE 215 AA; 25544 MW; 6824ED71729D1EC CMC64;
SC

Query Match 31.9%; Score 44; DB 16; Length 215;
Best Local Similarity 40.0%; Pred. No. 4.2;
Matches 8: Conservative 4; Mismatches 8; Indels 0; Gaps 0;
Db 8 GYANLAKAKIPMSIDITLD 27

RESULT 5
ID 091572 PRELIMINARY: PRT: 600 AA.
AC 091572
DC 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE CblAAC35225.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Brassicaceae; Brassicales; Brassicaceae; Arabidopsids.
NC [1]_TextID=3702;
RN [1]_TextID=3702;
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX Sato S., Nakamura Y., Kaneko T., Kato T., Asanizu E., Tabata S.;
AB "Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases."
PMBL: A9802528; A9804196.1.
DR InterPro: IPR000004; SAPB.
DR SMART: SM00118; SAPB: 1.
KT NON_TER 1 1
LN LINEAR POTENTIAL.
NM NUCLEONIC ACID SEQUENCE FROM N.A.
NC [1]_TextID=3702;
RN [1]_TextID=3702;
RP SEQUENCE 600 AA; 67338 MW; A78D9C89A82560D2 CMC64;
SC

Query Match 31.9%; Score 44; DB 10; Length 600;
Best Local Similarity 37.4%; Pred. No. 12;
Matches 9: Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Db 59 GLAVTLCTP---IASFCIAKLASTL 81

OY 2 F1ATLCTKXXXXVDFGID 20
 ID 419 F1ATLCTKXXXXVDFGID 437

RESULT 6

084690 PRELIMINARY: PRT: 607 AA.
 AC 084690:166 (TREMblrel. 01, Created)
 DT 01-JUN-1998 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE Structural viral protein.
 GN H GENE.
 OS Phocine distemper virus (PDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 RN 111111:TaxID=11240;
 RP SEQUENCE FROM N.A.
 RC STRAIN=PDV/DR8;
 RX MEDLINE=92113538; PubMed=1765768;
 RA Kovacs J., Blakenhorne-Moller M., Sharma B., Orvell C., North E.:
 "The nucleotide sequence and deduced amino acid composition of the
 haemagglutinin and fusion proteins of the morbillivirus phocid
 distemper virus." J. Gen. Virol. 72:2959-2966(1991).
 RL 121:Gen. Virol. 72:2959-2966(1991).
 RP SEQUENCE FROM N.A.
 RC STRAIN=PDV/DR8;
 RA Blakenhorne-Moller M.M.;
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 RI EMBL: Z69797; CA863428.1; --
 DR Pfam: PF00423; HN: 1; Hem-neuramids.
 SO SEQUENCE 607 AA; 16962 MW; BD9PB8DJA291C5B CRC64;

Query Match 31.9%: Score 44; DB 12; Length 607;

Best Local Similarity 28.6%: Pred. No. 12;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

OY 3 F1ATLCTKXXXXVDFGIDXXXXX1Q1 30
 ID 292 F1ATLCTKXXXXVDFGIDXXXXX1Q1 319

RESULT 7

056986 PRELIMINARY: PRT: 607 AA.
 AC 056986:198 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE H protein (Fragment).
 OS Phocine distemper virus (PDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 RN 111111:TaxID=11240;
 RP SEQUENCE FROM N.A.
 RC Vos H.W., Harder T.C., Kenler M., Osterhaus A.D.M.E.;
 RT "Expression of phocid distemper virus F and H genes in vaccinia
 virus".
 DL Submitted (FEB-1998) to the EMBL/Genbank/DBJ databases.
 RI EMBL: Z69797; CA863428.1; --
 DR InterPro: IPR000665; Hem-neuramids.
 DR Pfam: PF00423; HN: 1; Hem-neuramids.
 FT NON-UTR 607
 SO SEQUENCE 607 AA; 68842 MW; 622D6D5A8E179E5 CRC64;

Query Match 31.9%: Score 44; DB 12; Length 607;
 Best Local Similarity 28.6%: Pred. No. 12;

Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 OY 3 F1ATLCTKXXXXVDFGIDXXXXX1Q1 30
 ID 292 F1ATLCTKXXXXVDFGIDXXXXX1Q1 319

RESULT 8

08XN50 PRELIMINARY: PRT: 388 AA.
 AC 08XN50:202 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Capsular polysaccharide biosynthesis protein.
 GN CPE0488.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 RN 111111:TaxID=1502;
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 flesh-infecting bacterium." J. Bacteriol. 184:1996-2001(2002).
 RL EMBL: AP001187; BAB80194.1; --
 DR InterPro: IPR001295; Glycos.transf.1.
 DR Pfam: PF00534; Glycos.transf.1; 1.
 KW Complete proteome.
 SO SEQUENCE 388 AA; 45089 MW; 371FC2A13775F55A CRC64;

Query Match 31.5%: Score 42.5; DB 16; Length 388;

Best Local Similarity 33.3%: Pred. No. 5;
 Matches 11; Conservative 5; Mismatches 6; Indels 11; Gaps 1;

OY 1 GF1ATLCTKXXXXVDFGIDXXXXX1Q1LEPK 33
 ID 338 GF1A-----EPFIDISYRYKVLIDLR 359

RESULT 9

0985D7 PRELIMINARY: PRT: 807 AA.
 AC 0985D7:18 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE M7772A
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 RN 111111:TaxID=381;
 RP SEQUENCE FROM N.A.
 RC Kaneko T., Nakamura Y., Sato S., Asanuma E., Kato T., Sasamoto S.,
 RA Kaneko T., Nakamura Y., Sato S., Asanuma E., Kato T., Sasamoto S.,
 RA Watanabe A., Ideesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Naksumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Mesorhizobium loti: structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti." J. Bacteriol. 184:1996-2001(2002).
 RL DML Res. 7:331-338(2000).
 DR EMBL: AP003012; BAB54125.1; --
 DR InterPro: IPR000015; Flmb.usher.
 DR Pfam: PF00577; Usher; 1.
 KW Hypothetical protein; Complete proteome.

```

RT RT "Analysis of 74 kb of DNA located at the right end of the 330-kb
RT RT "Chlorella virus PBCV-1 genome."
RT RT "Virology 237:360-377(1997)."
RL RL (1)
RP RP "SEQUENCE FROM N.A." PubMed:10541099.
RP RP "Kaiser, A.D., Vollmert, M., Proll, D., Graves, M.V., Gurnon J.R., Xing W.,
RP RP "Liscac A.D., Nickerson, K.W., Van Etten J.L.;"
RP RP "Chlorella virus PBCV-1 encodes a functional homospemidline
RP RP "synthase."
RT RT "Virology 263:254-262(1999)."
RL RL (1)
RP RP "SEQUENCE FROM N.A." PubMed:11021991.
RP RP "Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;"
RT RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT RT "PBCV-1."
RL RL "Virology 276:27-36(2000)."
RP RP (14)
RP RP "SEQUENCE FROM N.A."
RP RP "Van Etten J.L.;"
RP RP "Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases."
RL RL (15)
RP RP "SEQUENCE FROM N.A."
RP RP "Van Etten J.L.;"
RP RP "Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases."
RL RL (16)
RP RP "SEQUENCE FROM N.A."
RP RP "Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases."
RL RL (17)
RP RP "SEQUENCE FROM N.A."
RP RP "Graves M.V., Van Etten J.L.;"
RP RP "Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases."
RL RL (18)
RP RP "SEQUENCE FROM N.A."
RP RP "Graves M.V., Van Etten J.L.;"
RP RP "Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases."
RL RL (19)
RP RP "SEQUENCE FROM N.A."
RP RP "Gurnon J.R., Graves M.V., Van Etten J.L.;"
RP RP "Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases."
RL RL "EMBL: U42580; AAC95945.1; ..."
RL RL "GenBank: AF060584; U083_MCP6-dh.G; ..."
RL RL "Pfam: PF03120; U083_MCP6-dh.G; ..."
RL RL "Pfam: PF03121; U083_MCP6-dh.N; ..."
RL RL "SEQUENCE 389 AA: 43458 MW: 8D5B5A98975426E0 CRC64:
SO SO

```


Db 176 IATLCATLEKXPAYRRCBYKDALLAOLIDPK 208
Search completed: May 9, 2003, 16:17:34
Job time : 29.3922 secs

09-MAY-2001; 2001WO-0540690.
 09-MAY-2000; 2000US-203063P.
 16-JUN-2000; 2000US-212042P.
 (GREN-) GREENVILLE HOSPITAL SYSTEM.
 Yu X, Wagner TE;
 WPI: 2002-154423/20.
 A procytotoxin comprising a cytotoxic peptide with at least one lysine residue which is converted into a cytotoxin in a target cell is used for treating cancer, particularly of the prostate, skin, ovary or lung

Disclousure; Page 12; 33pp; English.
 The present invention relates to a procytotoxin, comprising a cytotoxic peptide with at least one lysine residue bound by a peptide bond to at least one amino acid via the epsilon-amino group of the lysine residue. The procytotoxin is used to treat cancer, particularly of the prostate, ovary, lung or skin. Cytolytic peptide also known as channel-forming peptide, typically disrupt cell membrane, causing cell lysis and death (apoptosis) upon contact. The present sequence is procytotoxin.

Sequence 28 AA;
 Query Match 66.7%; Score 94; DB 23; Length 28;
 Best Local Similarity 71.0%; Pred. No. 9, 1e-09;
 Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

07 1 GIGAVLKXXXVTTGIPALISWIKXXXXXK 31
 07 1 GIGAVLK-----VTTGIPALISWIKRKKQ 27

Db 1 GIGAVLK-----VTTGIPALISWIKRKKQ 27

RESULT 2
 AAR34603 standard; peptide: 36 AA.
 AAR34603;
 09-AUG-1993 (first entry)
 Sequence of endosomolytic peptide.
 Endosomolytic peptide; endosome; transfection; gene transfer.
 Synthetic.
 WO9307283-A.
 15-APR-1993.
 28-SEP-1992; 92NO-EP02234.
 30-SEP-1991; 91US-0767788.
 30-SEP-1991; 91US-0768039.
 30-JAN-1992; 92US-0827103.
 30-JAN-1992; 92US-0827103.
 07-APR-1992; 92US-0864759.
 02-SEP-1992; 92US-0937788.
 (BOE) BOEHRINGER INGELHEIM INT GMBH.
 (GREN) GRENTECH INC.
 (UNNC-) UNIV NORTH CAROLINA.
 Birnsteiel ML, Colten M, Curiel D, Oberhauser B;
 Plank C, Schmidt MCM, Wagner E, Zalclokal K;
 WPI: 1993-134470/16.

Comps. for efficient transfection of higher eukaryotic cells -
 PT contains conjugate of nucleic acid, and affinity substance plus
 PT endosomolytic agent, esp. adenovirus, useful e.g. in gene
 transfer therapy
 Example; Page 196; 28pp; German.
 The peptide is endosomolytic. It can release the contents of
 endosomes into the cytoplasm. It can be covalently or non-
 covalently bound to a nucleic acid affinity substance (NAS) directly
 or via an NA-binding domain (NA - nucleic acid). The comps. are
 esp. used to introduce NA into human cells.
 Sequence 36 AA;
 Query Match 66.7%; Score 94; DB 14; Length 36;
 Best Local Similarity 71.0%; Pred. No. 1, 2e-08;
 Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

07 1 GIGAVLKXXXVTTGIPALISWIKXXXXXK 31
 07 1 GIGAVLK-----VTTGIPALISWIKRKKQ 27

Db 1 GIGAVLK-----VTTGIPALISWIKRKKQ 27

RESULT 3
 AAM01430 standard; peptide: 36 AA.
 AAM01430;
 27-JAN-1997 (first entry)
 Melittin peptide 1 with C-terminal oligo-Lys tail.
 Transfection; gene therapy; non-viral peptide; internalising factor:
 endosome; cytoplasm; oligo-lysine.
 Synthetic.
 US5447932-A.
 20-AUG-1996.
 30-SEP-1991; 91US-0767788.
 23-SEP-1992; 92US-0948357.
 30-SEP-1991; 91US-0767788.
 30-SEP-1991; 91US-0768039.
 30-JAN-1992; 92US-0827103.
 30-JAN-1992; 92US-0827103.
 07-APR-1992; 92US-0864759.
 02-SEP-1992; 92US-0937788.
 (BOE) BOEHRINGER INGELHEIM INT GMBH.
 (GREN) GRENTECH INC.
 Birnsteiel ML, Colten M, Curiel DT, Oberhauser B;
 Plank C, Schmidt MCM, Wagner E, Zalclokal K;
 WPI: 1996-392650/39.
 Introduction of nucleic acid into higher eukaryotic cells - as
 PT complex with nucleic acid affinity agent and endosomolytic agent
 Example 37; Column 95-96; 109pp; English.
 A nucleic acid can be introduced into the cytoplasm of a higher
 eukaryotic cell by contacting the cell with a compo. comprising a
 complex, where the complex comprises a nucleic acid of interest, a
 substance with affinity for the nucleic acid and an endosomolytic
 agent, which allows internalisation of the complex into the cell by
 endocytosis, and is released into the cytoplasm by endosomolysis.
 The present sequence is a melittin peptide designated mel 1, which

QY 1 GIGAVLKXXXXVLTGCPALISWIKXXXXXK 31
 ID AAV77660 standard; peptide: 36 AA.
 NC AAV77660:
 CC 12-MAY-2000 (first entry)
 DE Melittin peptide mel 1 with oligolysine tail.
 XX Endosomalytic; transfection; tumour; cancer; vaccine; interleukin-2;
 XX Immune modulation; gene therapy; cystic fibrosis; lysosomotropic;
 XX melittin; antimicrobial; respiratory.
 CS Synthetic.
 XX US6022735-A.
 PN 08-FEB-2000.
 PD 25-MAY-1995. 95US-0449741.
 XX 23-SEP-1992. 92US-0948357.
 PR 30-SEP-1991. 91US-0767786.
 PR 30-SEP-1991. 91US-0768039.
 PR 30-JAN-1992. 92US-0827102.
 PR 30-JAN-1992. 92US-0827103.
 PR 07-APR-1992. 92US-0864759.
 PR 02-SEP-1992. 92US-0937788.
 XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 PA (UNCL-) UNIT NORTH CAROLINA.
 XX Wagner E, Plank C, Birnstiel ML, Schmidt WM, Curtel DT;
 PI Zalcoual K, Oberhauser B, Cotten M;
 DR WPI: 2000-160379/14.
 XX Transfection of nucleic acid complexes into eukaryotic cells by
 PT compositions containing endosomalytic viral peptide, useful in somatic
 PT gene therapy and in the treatment of cystic fibrosis.
 XX Example 37; column 74, 93pp; English.
 CC The invention relates to the use of an endosomalytic viral peptide in a
 CC composition for the transfection of a high eukaryotic cell with a nucleic
 CC acid complex. The peptide has the ability of being internalized into the
 CC higher eukaryotic cell, either on its own or as a component of the
 CC nucleic acid complex, where it has the ability to release contents of the
 CC complex into the cytoplasm. The complex then enters the
 CC higher eukaryotic cell into the cytoplasm. The complex enters the
 CC cell on tumour cells for use as a cancer vaccine, where the DNA to be
 CC introduced into the tumour cells codes for an immune modulating substance
 CC such as interleukin-2. The composition is also useful in somatic gene
 CC therapy and in the treatment of cystic fibrosis. Use of the endosomalytic
 CC peptide with the conjugate-DNA complex increases the efficiency of the
 CC transfection system much better than the use of the lysosomotropic
 CC peptide. The present sequence represents a melittin peptide
 CC with an oligolysine tail, synthesised for use in transfection of
 CC hepatocytes.
 XX Sequence 36 AA:
 SQ Query Match 66.7%: Score 94: DB 21: Length 36:
 Best Local Similarity 71.0%: Pred. No. 1,2e-08:
 Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:

QY 1 GIGAVLKXXXXVLTGCPALISWIKXXXXXK 31
 ID AAV77660 standard; peptide: 36 AA.
 NC AAV77660:
 CC 12-MAY-2000 (first entry)
 DE Melittin peptide mel 1 with oligolysine tail.
 XX Endosomalytic; transfection; tumour; cancer; vaccine; interleukin-2;
 XX Immune modulation; gene therapy; cystic fibrosis; lysosomotropic;
 XX melittin; antimicrobial; respiratory.
 CS Synthetic.
 XX US6022735-A.
 PN 08-FEB-2000.
 PD 25-MAY-1995. 95US-0449741.
 XX 23-SEP-1992. 92US-0948357.
 PR 30-SEP-1991. 91US-0767786.
 PR 30-SEP-1991. 91US-0768039.
 PR 30-JAN-1992. 92US-0827102.
 PR 30-JAN-1992. 92US-0827103.
 PR 07-APR-1992. 92US-0864759.
 PR 02-SEP-1992. 92US-0937788.
 XX (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 PA (UNCL-) UNIT NORTH CAROLINA.
 XX Wagner E, Plank C, Birnstiel ML, Schmidt WM, Curtel DT;
 PI Zalcoual K, Oberhauser B, Cotten M;
 DR WPI: 2000-160379/14.
 XX Transfection of nucleic acid complexes into eukaryotic cells by
 PT compositions containing endosomalytic viral peptide, useful in somatic
 PT gene therapy and in the treatment of cystic fibrosis.
 XX Example 37; column 74, 93pp; English.
 CC The invention relates to the use of an endosomalytic viral peptide in a
 CC composition for the transfection of a high eukaryotic cell with a nucleic
 CC acid complex. The peptide has the ability of being internalized into the
 CC higher eukaryotic cell, either on its own or as a component of the
 CC nucleic acid complex, where it has the ability to release contents of the
 CC complex into the cytoplasm. The complex then enters the
 CC higher eukaryotic cell into the cytoplasm. The complex enters the
 CC cell on tumour cells for use as a cancer vaccine, where the DNA to be
 CC introduced into the tumour cells codes for an immune modulating substance
 CC such as interleukin-2. The composition is also useful in somatic gene
 CC therapy and in the treatment of cystic fibrosis. Use of the endosomalytic
 CC peptide with the conjugate-DNA complex increases the efficiency of the
 CC transfection system much better than the use of the lysosomotropic
 CC peptide. The present sequence represents a melittin peptide
 CC with an oligolysine tail, synthesised for use in transfection of
 CC hepatocytes.
 XX Sequence 36 AA:
 SQ Query Match 66.7%: Score 94: DB 22: Length 36:
 Best Local Similarity 71.0%: Pred. No. 1,2e-08:
 Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:

Best Local Similarity 71.0%; Pred. No. 1.2e-08;
Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;
OY 1 GIGAVLXXXXXXVTTGTPALISWIKXXXXXK 31
Db 1 GIGAVLK----VLTGTPALISWIKRRROK 27

RESULT 8

AAAP91340 ID AAP91340 standard; peptide: 26 AA.

XX AAP91340;

DT 19-MAR-1990 (first entry)

DE Amino acid sequence of Melittin.

KM Melittin; lytic peptide; antimicrobial peptide; disease-resistant

XX Lirichophyce.

PN M06904371-A.

XX 18-MAY-1989.

PE 02-NOV-1988; 88MO-US03908.

XX 02-NOV-1987; 87US-0115941.

PA (LOUUI) LOUISIANA STATE UNIV.

XX Jaynes JM, Derrick KS;

DR WPI: 1989-165650/22.

XX Transformed plants contg. heterologous gene - expressing antimicrobial

CC agent, or polypeptide high in essential amino acids

XX Table 1; : 56pp; English.

PS Amino acid sequence of Melittin as an exemplary lytic peptide for

CC use as an antimicrobial peptide contemplated for use in plant

CC agents and organisms in the invention. Melittins are generally

CC less preferred in the invention because of their lack of specificity for

CC bacteria and fungi.

XX Sequence 26 AA:

OY 1 GIGAVLXXXXXXVTTGTPALISWIKXXXXXK 38

Db 1 GIGAVLK----VLTGTPALISWIK----RK----RQO 26

RESULT 9

AAAR13908 ID AAR13908 standard; peptide: 26 AA.

XX AAR13908;

DT 27-NOV-1991 (first entry)

DE Guanidinated melittin.

XX B-cell; T-cell; epitope; carrier; hapten.

XX Synthetic.

FT Key Modified-site 7 Location/Qualifiers

FT Modified-site /label= guanidinated
FT 21 /label= guanidinated
FT Modified-site /label= guanidinated
FT 23 /label= guanidinated
FT Modified-site /label= laccopyranoside moiety.
FT 25 /label= hinge
FT Region 12..14
FT Binding-site 7..19
FT Binding-site /label= T cell epitope
FT 20-26 /label= B cell epitope

XX M09110676-A.

XX 25-JUL-1991.

PD 19-JAN-1991; 91MO-US00413.

XX 15-JAN-1991; 91US-0641459.

XX 19-JUN-1990; 90US-0467386.

XX (UYRO-) ROCKEFELLER UNIV.

XX Fehner PF, King TP, Kochumian L;

XX WPI: 1991-281140/38.

XX Immunogenic mole. contg. a T-cell epitope - comprising an

XX amphiphilic peptide carrier bound to a hapten which is a B-cell

XX epitope.

XX Claim 7; Fig 1; 36pp; English.

PS Melittin is a peptide isolated from bee venom. Analogues can be

CC used as an antimicrobial peptide contemplated for use in plant

CC agents and organisms in the invention. The epsilon-amino

CC analogues must contain at least one T cell epitope (see features)

CC and may additionally contain a B cell epitope. The epsilon-amino

CC cps. of the lysine residues at the 7, 21 and 23 positions are

CC converted to guanidino gps and that at the C-terminal is blocked.

CC The analogues may be truncated from the N-terminal to remove the

CC hydrophilic residues 20-26.

XX Sequence 26 AA:

OY 1 GIGAVLXXXXXXVTTGTPALISWIKXXXXXK 38

Db 1 GIGAVLK----VLTGTPALISWIK----RK----RQO 26

RESULT 10

AAAR45114 ID AAR45114 standard; peptide: 26 AA.

XX AAR45114;

DT 28-JUN-1994 (first entry)

DE Melittin for N-terminal lipophilic substitution.

XX Ion channel: magainin, pGLA, XRF, GFP; cecropin, sarcocoxin;

XX inhibition: target cells; virus; vitally-infected cells;

XX antimicrobial; antiviral; antitumour; antiparasitic;

XX spermidic; wound healing; burn; infection.

XX Synthetic.

PN M09324138-A.
 XX 09-DEC-1993.
 PF 27-MAY-1993: 93MO-US05192.
 XX
 PR 01-JUN-1992: 92US-0891201.
 XX
 PA (MAGA-) MAGALININ PHARM INC.
 XX
 XX Karl U.
 XX
 DR WPI: 1993-405419/50.
 XX
 PF Peptide(s) of proteins with an N-terminal lipophilic substit. -
 PT used for inhibiting growth of target cell, virus or
 PT virally-infected cell
 XX
 XX Disclosure: Page 39: 113pp: English.
 XX
 CC A novel compen. for inhibiting growth of a target cell, virus or
 CC virally-infected cell comprises a peptide of formula T-(M)-X (I).
 CC X is a biologically active amphiphilic ion channel-forming peptide
 CC or protein; pref. a magalinin peptide, a pica peptide, a xpf
 CC peptide, a cfp peptide, a cecropin or a sarcotoxin.
 CC T is a cationic or the N-terminal amino group.
 CC M is a lipophilic or aromatic or alkylaromatic).
 CC hydrocarbon (alkyl or aromatic or alkylaromatic).
 CC T is pref. an octenoyl group.
 CC W is T or hydrogen.
 CC
 CC Melittin is an amphiphilic peptide and an example of X.
 CC The N-terminal subseq. peptides and proteins have increased
 CC biological activity as compared with unsubseq. peptides or proteins
 CC which are biologically active as compared with unsubseq. peptides or
 CC proteins. They can be used as antimicrobial agents, antiviral agents,
 CC antitumour agents, antiparasitic agents or spermicides and
 CC can also exhibit other bioactive functions. They can also be
 CC used in promoting or stimulating wound healing, for the treatment
 CC of external burns and to treat and/or prevent skin and burn
 CC infections or eye infections.
 XX
 SO Sequence 26 AA:
 Query Match 66.0%: Score 93: DB 14: Length 26:
 Best Local Similarity 68.4%: Pred. No. 1,2e-08:
 Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:
 OY 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 38
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 RESULT 11
 AAR35383 standard: peptide: 26 AA.
 ID AAR35383 standard: peptide: 26 AA.
 AC AAR35383:
 XX
 DT 07-JUN-1993 (first entry)
 XX
 XX Melittin used to treat oral infections.
 DE
 XX Adverse oral conditions: amphiphilic; anti-bacterial; anti-viral;
 KM anti-fungal; dental plaque; dental caries; periodontal disease;
 KM gingivitis; leishmaniasis; ion-channel forming; honeybee; venom.
 OS
 XX Apis mellifera.
 XX
 PN M09301723-A.
 XX
 XX 04-FEB-1993.
 XX
 XX 09-JUL-1992: 92MO-US05757.
 PF

XX 25-JUL-1991: 91US-0735070.
 XX
 XX (MAGA-) MAGALININ PHARM INC.
 XX
 XX Berkowitz B, Jacob L.
 PF
 DR WPI: 1993-058434/07.
 XX
 PF Peptide(s) for prophylaxis and treatment of oral disorders - used
 PT for periodontal disease, plaque, dental caries, gingivitis, etc.
 XX
 PS Claim 2: Page 130: 143pp: English.
 XX
 CC Melittin is an amphiphilic peptide which is isolated from honeybee
 CC venom. The peptide is known to be cytolytic (see Hoppe-Seyler's
 CC Zeitschrift Physiol. Chem. 348: 37-50, 1987) and is a preferred
 CC peptide for use in preventing or treating adverse oral conditions.
 CC The peptide is suitable for use in oral compositions to treat oral
 CC conditions such as gingivitis, periodontal disease, dental caries
 CC and gingivitis. The anti-bacterial action will also be useful against
 CC bacteria associated with dental implant infections and the peptides
 CC can stimulate the healing of wounds in the oral cavity.
 XX
 SO Sequence 26 AA:
 Query Match 66.0%: Score 93: DB 14: Length 26:
 Best Local Similarity 68.4%: Pred. No. 1,2e-08:
 Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:
 OY 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 38
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 DB 1 GIGAVLKXXXVVTGCPALISWIKXXXXXKXXXXXKQ 26
 RESULT 12
 AAR39759 standard: peptide: 26 AA.
 ID AAR39759 standard: peptide: 26 AA.
 AC AAR39759:
 XX
 DT 08-FEB-1994 (first entry)
 XX
 XX Melittin peptide.
 DE
 XX Melittin; amphiphilic; honeybee venom; cytolytic; antifungal;
 KM antibiotic; antiparasitic; antimicrobial; spermicidal; antiviral;
 KM
 XX Apis mellifera.
 XX
 PN US235038-A.
 XX
 PD 10-AUG-1993.
 XX
 XX 22-JAN-1991: 91US-0643343.
 XX
 XX 22-JAN-1991: 91US-0643343.
 XX
 XX (TORR-) TORR PINES INSR MOLECULAR STUDIES.
 XX
 XX Blondelle SE, Houghten RA:
 PF
 DR WPI: 1993-264689/33.
 XX
 PF Peptide and substitution peptide analogues of melittin - used as
 PT antibiotics; antifungal agents; antitumour agents; antiviral
 PT agents or to stimulate wound healing
 XX
 PS Disclosure: Page 1: 36pp: English.
 XX
 CC The peptide melittin (AAR39759) from honeybee venom was used to
 CC produce deletion analogues (AAR39759-R39783) and substitution
 CC analogues (AAR39784-R39800) (AAR39827-R39839). The peptides can be used

CC to inhibit growth of the target cell, thus exhibiting antimicrobial,
CC anti-viral, antibiotic, anti-tumor, antiparasitic, antifungal and
CC spermicidal properties.

XX Sequence 26 AA:

Query Match 66.0% Score 93: DB 14: Length 26:
Best Local Similarity 68.4% Pred. No. 1.2e-08:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGAVLKXXXXVTGTGPAISWIKXXXXXRRQ 38
DB 1 GIGAVLK----VLTGCPALISWIK----RK----RQ 26

RESULT 13

AA055989 standard: peptide: 26 AA.

XX AA055989:

DT 19-DEC-1994 (first entry)

DE Ion channel forming peptide (mellitin):

KM Ion channel forming peptide: tumour; skin disease; malignancy;
KM melanoma; carcinoma; basal cell; squamous cell; magalain; PGLa;
KM CFP peptides: sarcotoxin; sarcotoxin; mellitin; apidocin; defensins;
KM basic protein; bacteria permeability increasing protein;
KM peflorin.

OS Apis mellifera.

PN W09412206-A.

XX 09-JUN-1994.

PF 03-DEC-1993: 93MO-US11885.

PR 03-DEC-1992: 92US-0984957.

PA (MAGA-) MAGALIN PHARM INC.

XX Jacob LS, Maloy WL.

DR WPI: 1994-19965/24.

PT Treating skin cancer with ion channel forming peptide(s) - e.g.
PT magalain, mellitin etc., specifically for treating melanoma

XX Disclosure: Page 122: 136pp: English.

CC The peptide is used to treat dermatological malignancies. It
CC may be used to treat especially melanoma but also basal cell and
CC squamous cell carcinomas. It can be used together with an ion which
CC also inhibits/prevents growth of the target cell. Peptides used for
CC such therapy include magalain, PGLa or CFP peptides; sarcotoxins,
CC sarcotoxins, mellitin, apidocin, defensins, major basic protein of
CC eosinophils, bactericidal permeability increasing protein and peflorin.
CC See also AM055989-055997.

XX Sequence 26 AA:

Query Match 66.0% Score 93: DB 15: Length 26:
Best Local Similarity 68.4% Pred. No. 1.2e-08:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGAVLKXXXXVTGTGPAISWIKXXXXXRRQ 38
DB 1 GIGAVLK----VLTGCPALISWIK----RK----RQ 26

RESULT 14

AA050565 standard: peptide: 26 AA.

XX AA050565:

DT 18-OCT-1994 (first entry)

DE Amphiphilic peptide #114.

KM Amphiphilic: ion forming; gynaecological malignancy; magalain; PGLa;
KM xPF; CFP; sarcotoxin; mellitin; apidocin; defensins;
KM major basic protein; eosinophils; uterine; cervical; cancer;
KM bacterial permeability increasing protein; ovarian; stage IC.
OS Synthetic.

PN W09405313-A.

PD 17-MAR-1994.

XX 16-AUG-1993: 93MO-US07798.

PR 31-AUG-1992: 92US-0937462.

XX (MAGA-) MAGALIN PHARM INC.

PA Baker MA, Jacob LS, Maloy WL.

XX WPI: 1994-100851/12.

PT Treating gynaecological tumours with amphiphilic peptide(s) -
PT which form ion channels, e.g. magalain or PGLa peptide(s),
PT partic. for treating ovarian, uterine or cervical cancers.

XX Disclosure: Page 117: 130pp: English.

CC The sequences given in AA050452-568 represent amphiphilic, ion forming
CC peptides which may be used to treat gynaecological malignancy.
CC These peptides are based on magalain, PGLa, xPF or CFP, a sarcotoxin, a
CC sarcotoxin, mellitin, an apidocin, a defensin, major basic protein of
CC eosinophils or a bacterial permeability increasing protein. These
CC peptides are esp. used to treat ovarian, esp. stage IC, uterine or
CC cervical cancers.

XX Sequence 26 AA:

XX Query Match 66.0% Score 93: DB 15: Length 26:

XX Best Local Similarity 68.4% Pred. No. 1.2e-08:

XX Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGAVLKXXXXVTGTGPAISWIKXXXXXRRQ 38

DB 1 GIGAVLK----VLTGCPALISWIK----RK----RQ 26

RESULT 15

AA050430 standard: peptide: 26 AA.

XX AA050430:

DT 17-OCT-1994 (first entry)

DE Amphiphilic peptide #95, mellitin derivative.

KM Amphiphilic peptide; aprotic organic solvent; alcohol; antitumour;
KM antileukemic; antimicrobial; antitumour; antiparasitic; anticancer;
KM antidiabetic; human; animal; plant; ion-channel; forming peptide.
OS Apis mellifera.

PN W09405308-A.

PD 17-MAR-1994.
 XX 13-AUG-1993: 93MO-US07694.
 XX
 XX 28-AUG-1992: 92US-0936504.
 PR
 XX (MAGA-) MACAINTIN PHARM INC.
 PA
 XX
 XX W1111ems J1:
 PT WPI: 1994-100846/12.
 XX
 XX
 PT Purifying amphiphilic protein or peptide by solvent extr.
 PT partic. for recombinant, ion-channel forming peptide(s) such as
 PT magainins, avoids use of chaotropic agents.
 XX
 XX
 PS Disclosure: Page 117: 135pp: English.
 XX
 XX The sequences given in AA05036-451 are amphiphilic peptides which
 CC were isolated by the method of the invention. A material containing
 CC amphiphilic peptides such as these, was treated with a mixture of
 CC aprotic organic solvent and alcohol to form a single miscible
 CC solution. This solution was then treated with a aqueous solution to
 CC form an aqueous phase solution containing the peptides and an
 CC organic solvent phase, and the peptides were isolated from the
 CC aqueous phase. The isolated peptides may be used as antibiotic,
 CC antitumor, antifungal, antiviral, antiparasitic, antipneumonia,
 CC and/or antiviral agents for treatment of humans, animals or plants.
 CC These peptides are esp. ion-channel forming peptides which enable
 CC biologically active ions to enter cells.
 XX

XX Sequence 26 AA:

Query Match 66.0% Score 93: DB 15: Length 26:
 Best Local Similarity 68.4% Pred. No. 1.2e-06:
 Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:
 QY 1 GIGAVLKXXXXVTTGLPALISWIKXXXXRXXXXRQ 38
 ||||||| ||||||||| || |||
 DB 1 GIGAVLK----VTTGLPALISWIK---RK---RQ 26

Search completed: May 9, 2003, 16:15:35
 Job time : 54.6471 secs

GenCore version 5.1.4-p5 4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OK protein - protein search, using sw model

Run on: May 9, 2003, 16:13:10 ; Search time 13.7843 seconds
(without alignments)
81.112 Million cell updates/sec

Title: US-09-851-422A-2

Sequence: 1 GIGAVLXXXXXVLTGTLPALISWIKXXXXXRRXXXXXHQ 38
Perfect score: 141

Scoring table: BLOSUM62
Gapop 10.0 / Gapext 0.5

Searched: 262574 seqs; 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing list 45 summaries

Database :

- 1: /cgn2-6/podata/1/1aa/5A.COMB.pep.*
- 2: /cgn2-6/podata/1/1aa/5B.COMB.pep.*
- 3: /cgn2-6/podata/1/1aa/6A.COMB.pep.*
- 4: /cgn2-6/podata/1/1aa/6B.COMB.pep.*
- 5: /cgn2-6/podata/1/1aa/6C.COMB.pep.*
- 6: /cgn2-6/podata/1/1aa/6D.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	66.7	36	1	US-07-948-357-12
2	94	66.7	36	1	US-08-450-417-12
3	94	66.7	36	3	US-08-449-741-12
4	94	66.7	36	3	US-08-449-741-12
5	94	66.7	36	4	US-07-643-343A-12
6	93	66.0	26	1	US-07-643-343A-12
7	93	66.0	26	1	US-07-626-589-3
8	93	66.0	26	1	US-08-079-445-7
9	93	66.0	26	1	US-07-943-709-58
10	93	66.0	26	1	US-08-236-410-3
11	93	66.0	26	1	US-08-218-608-8
12	93	66.0	26	1	US-08-434-120-114
13	93	66.0	26	1	US-08-465-421-3
14	93	66.0	26	1	US-08-465-421-3
15	93	66.0	26	1	US-08-984-284-19
16	93	66.0	26	1	US-08-984-284-19
17	93	66.0	26	1	US-08-984-284-19
18	93	66.0	26	1	US-08-984-284-19
19	93	66.0	26	1	US-08-984-284-19
20	93	66.0	26	1	US-08-984-284-19
21	93	66.0	26	1	US-08-984-284-19
22	93	66.0	26	1	US-08-984-284-19
23	93	66.0	26	1	US-08-984-284-19
24	93	66.0	26	1	US-08-984-284-19
25	93	66.0	26	1	US-08-984-284-19
26	93	66.0	26	1	US-08-984-284-19
27	93	66.0	26	1	US-08-984-284-19

28	93	66.0	26	4	US-09-319-730-13	Sequence 13, Appl
29	93	66.0	26	4	US-09-319-730-13	Sequence 93, Appl
30	93	66.0	26	4	US-08-842-322-28	Sequence 28, Appl
31	93	66.0	26	4	US-08-842-322-28	Sequence 10, Appl
32	93	66.0	26	4	US-08-842-322-28	Sequence 11, Appl
33	92	65.2	24	2	US-08-553-221-2	Sequence 2, Appl
34	92	65.2	24	2	US-08-553-221-2	Sequence 23, Appl
35	92	65.2	25	1	US-07-643-343A-22	Sequence 22, Appl
36	92	65.2	25	1	US-07-643-343A-22	Sequence 24, Appl
37	92	65.2	25	1	US-07-643-343A-24	Sequence 25, Appl
38	92	65.2	25	1	US-07-643-343A-25	Sequence 11, Appl
39	92	65.2	25	1	US-07-643-343A-41	Sequence 41, Appl
40	92	65.2	25	1	US-07-643-343A-41	Sequence 43, Appl
41	92	65.2	26	1	US-07-643-343A-43	Sequence 44, Appl
42	92	65.2	26	1	US-07-643-343A-45	Sequence 45, Appl
43	92	65.2	26	1	US-07-643-343A-45	Sequence 7, Appl
44	92	65.2	26	1	US-08-141-324-7	Sequence 1, Appl
45	92	65.2	26	1	US-08-295-086-1	

ALIGNMENTS

RESULT 1
US-07-948-357-12
Sequence 12, Application US/07948357
Patent No. 5547932
GENERAL INFORMATION:
APPLICANT: David T.
APPLICANT: Birnstiel Max L.
APPLICANT: Colten, Matthew
APPLICANT: Wagner, Ernest
APPLICANT: Zalloukai, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Bernd
APPLICANT: Schmitt, Peter
TITLE OF INVENTION: Composition for Introducing Nucleic Acid
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: Washington
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Present in Release #1.0, Version #1.25
CURRICULUM VITAE: US/07/948-357
APPLICATION NUMBER: US/07/948-357
FILING DATE: 19920923
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652, 0940004
TELEPHONE: (202) 833-8716
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: AMINO ACID
STRATEGY: single
TOPOLGY: both
MOLECULE TYPE: peptide
US-07-948-357-12
Query Match 66.7% Score 94; DB 1; Length 36;
Best Local Similarity 71.0%; Pred. No. 5.9e-10;

Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:
OY 1 GIGAVLXXXXXVITGCPALISMIXKXXXXR 31
DB 1 GIGAVLX-----VITGCPALISMIXKXXXXR 27

RESULT 2

US-08-450-417-12

Sequence 12, Application US/08450417
Patent No. 6027735
GENERAL INFORMATION:
APPLICANT: Curitel, David T.
APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernest
APPLICANT: Zalloukai, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
TITLE OF INVENTION: Composition for introducing Nucleic Acid
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
STATE: Washington
COUNTRY: U.S.A.
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATOR: IBM PC compatible
SOFTWARE: PC DOS/MS-DOS
CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.25
APPLICATION NUMBER: US/08/450,417
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/948,357
PUBLICATION NUMBER: 19990213
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652, 0940004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 456-0800
FAX: (202) 456-0800
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-450-417-12

Query Match 66.7% Score 94: DB 2: Length 36:
Best Local Similarity 71.0% Pred. No. 5.9e-10:
Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:

OY 1 GIGAVLXXXXXVITGCPALISMIXKXXXXR 31
DB 1 GIGAVLX-----VITGCPALISMIXKXXXXR 27

RESULT 3

US-08-449-741-12

Sequence 12, Application US/08449741
Patent No. 6027735
GENERAL INFORMATION:
APPLICANT: Curitel, David T.

APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernest
APPLICANT: Zalloukai, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
TITLE OF INVENTION: Composition for introducing Nucleic Acid
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, Suite 600
STATE: D.C.
CITY: Washington
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATOR: IBM PC compatible
SOFTWARE: PC DOS/MS-DOS
CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.25
APPLICATION NUMBER: US/08/449,741
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 34,131
REFERENCE/DOCKET NUMBER: 0652, 0940007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
FAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-449-741-12

Query Match 66.7% Score 94: DB 3: Length 36:
Best Local Similarity 71.0% Pred. No. 5.9e-10:
Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:

OY 1 GIGAVLXXXXXVITGCPALISMIXKXXXXR 31
DB 1 GIGAVLX-----VITGCPALISMIXKXXXXR 27

RESULT 4

US-08-449-754-12

Sequence 12, Application US/08449754
Patent No. 6077663
GENERAL INFORMATION:
APPLICANT: Curitel, David T.
APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernest
APPLICANT: Zalloukai, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
TITLE OF INVENTION: Composition for introducing Nucleic Acid
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
STATE: D.C.
CITY: Washington
COUNTRY: U.S.A.

ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/449,754
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/946,357
 FILING DATE: 1997-09-23
 ATTORNEY/AGENT INFORMATION:
 NAME: Fesmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652.0940004
 TELEPHONE: (202) 466-0800
 TELEFAX: (202) 466-0816
 INFORMATION FOR SEO ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 36 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: both
 MOLECULE TYPE: peptide
 US-08-449,754-12

Query Match 66.7% Score 94: DB 3: Length 36:
 Best Local Similarity 71.0% Pred. No. 5,9e-10:
 Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:
 Oy 1 GIGAVLKKXXVVTGGLPALISWIKKXXRK 31
 Db 1 GIGAVLK-----VLTGGLPALISWIKRKROK 27

RESULT 5
 US-09-465-646-12
 Sequence 12: Application US/09465646
 Patent No. 6294322
 GENERAL INFORMATION:
 APPLICANT: Curitel, David T.
 APPLICANT: Birnstiel, Max L.
 APPLICANT: Cotten, Matthew
 APPLICANT: Wagner, Ernst
 APPLICANT: Zalioukai, Kurt
 APPLICANT: Plank, Christian
 APPLICANT: Oberhauser, Berndt
 APPLICANT: Schlegel, Stefan
 TITLE OF INVENTION: COMPOSITION FOR IMPROVING NUCLEIC ACID
 TITLE OF INVENTION: COMPLEXES INTO HIGHER EDUCATION CELLS
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Ave., N.W., Suite 600
 STATE: Washington
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/465,646
 FILING DATE: HERMITH
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/449,754
 FILING DATE: May 25, 1995

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/937,788
 FILING DATE: September 2, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/864,759
 FILING DATE: April 7, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/827,102
 FILING DATE: January 30, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/827,103
 FILING DATE: January 30, 1992
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/67,788
 FILING DATE: September 30, 1991
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/768,039
 FILING DATE: September 30, 1991
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Karmage, Andrea J.
 REGISTRATION NUMBER: 33,703
 REFERENCE/DOCKET NUMBER: 0652.0940009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-6566
 INFORMATION FOR SEO ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 36 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: both
 MOLECULE TYPE: peptide
 US-09-465-646-12

Query Match 66.7% Score 94: DB 4: Length 36:
 Best Local Similarity 71.0% Pred. No. 5,9e-10:
 Matches 22: Conservative 1: Mismatches 4: Indels 4: Gaps 1:
 Oy 1 GIGAVLKKXXVVTGGLPALISWIKKXXRK 31
 Db 1 GIGAVLK-----VLTGGLPALISWIKRKROK 27

RESULT 6
 US-07-643-343a-1
 Sequence 1: Application US/0764343A
 Patent No. 5250038
 GENERAL INFORMATION:
 APPLICANT: Boudhien, Sylvie E.
 APPLICANT: Houghlen, Richard A.
 TITLE OF INVENTION: Deletion and Substitution
 NUMBER OF SEQUENCES: 45
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Carelia, Byrne, Bain, Giffilian,
 STREET: 6 Becker Farm Road
 CITY: Roseland
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch diskette
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: DNA V2

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/643,343A
FILING DATE: 1994-01-22
CITING DATE: 1994-01-22
CLASSIFICATION: 510
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M., 025
REGISTRATION NUMBER: 24,025
TELEPHONE: 201-994-1700
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORGANISM: Apis mellifera
FEATURE:
NAME/KEY: melittin peptide
PARENT INFORMATION:
AUTHORS: Jentsch, J.
TITLE: Sequenzanalyse des Melittins aus
TITLE: den tryptischen und peptischen
TITLE: Spaltst. chem.
JOURNAL: Hoppe-Seyler's Zeitschrift
VOLUME: 36
PAGES: 37-50
DATE: 1987
US-07-643-343A-1

Query Match 66.0% Score 93; DB 1; Length 26;
Best Local Similarity 68.4% Pred. No. 6,1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Cy 1 GIGAVLXXXXVLTGCPALISWIKXXXXXHQ 38
Db 1 GIGAVLK-----VLTGCPALISWIK----RK----NQ 26

RESULT 7
US-07-626-589-3
Sequence 3; Application US/07/626589
Patent No. 5254698
GENERAL INFORMATION:
APPLICANT: Bomalaski, John S.
APPLICANT: Clark, Michael A.
INVENTOR: Short, Robert
TITLE OF INVENTION: INHIBITORY PHOSPHOLIPASE ACTIVATING PROTEIN
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: United States of America
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
FILING DATE: 05/07/626,589
APPLICATION NUMBER: 05/07/626,589
FILING DATE: 19901206

```

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hohenschuetz, Liza D.
REGISTRATION NUMBER: 23,712
TELEPHONE: (219) 568-3100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (219) 568-3100
TELEFAX: (219) 568-3439
INFORMATION FOR SEQ ID NO: 3;
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-07-626-589-3

Query Match 66.0% Score 93; DB 1; Length 26;
Best Local Similarity 68.4% Pred. No. 6,1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Cy 1 GIGAVLXXXXVLTGCPALISWIKXXXXXHQ 38
Db 1 GIGAVLK-----VLTGCPALISWIK----RK----NQ 26

RESULT 8
US-08-079-445-7
Sequence 7; Application US/08079445
Patent No. 5440016
GENERAL INFORMATION:
APPLICANT: Biordelle, Sylvie E.
APPLICANT: Phyllis, Clemencia
APPLICANT: Elchler, Julia
INVENTOR: Biordelle, Sylvie E.
TITLE OF INVENTION: PEPTIDES HAVING ANTI-MICROBIAL
TITLE OF INVENTION: HEMOLYTIC, ANTI-MELITTIN OR ANTI-ENZYMATIC ACTIVITY
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
FILING DATE: 05/08/079,445
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathlyn A.
REGISTRATION NUMBER: 11,815
REFERENCE/DOCKET NUMBER: HP 9648
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7;
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
STRANDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 26
OTHER INFORMATION: /noce="C-terminal amino acid 1s
OTHER INFORMATION: amidated"
US-08-079-445-7

```


CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/218, 608
 FILING DATE: 28-MAR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: GATES, EDWARD R.
 REGISTRATION NUMBER: 31,616
 REFERENCE/DOCKET NUMBER: M0656/7013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-720-3500
 TELEFAX: 617-720-3411
 INTRODUCTION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 26 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYDROLYZABLE: NO
 ARTIFICIAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Aphis mellifera
 US-08-218-608-8

Query Match 66.0%; Score 93; DB 1; Length 26;
 Best Local Similarity 68.4%; Pred. No. 6.1e-10;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
 Oy 1 GIGAVLXXXXXVLTGTPALISMIXXXXXXXXRQ 38
 Db 1 GIGAVLX---VLTGTPALISMIX---RK---RQ 26

RESULT 12

US-08-434-120-114
 Sequence 114, Application US/08434120
 Patent No. 5635479
 GENERAL INFORMATION:
 APPLICANT: Baker, Margaret A.
 APPLICANT: Jacob, Leonard S.
 TITLE OF INVENTION: Method of Synecological
 TITLE OF INVENTION: Biologically Active Peptides
 TITLE OF INVENTION: Biologically Active Peptides
 NUMBER OF SEQUENCES: 117
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Carrelle, Byrne, Bain, Gillfillan,
 ADDRESSEE: Cecchi & Stewart
 STREET: Becker Farm Road
 CITY: Roseland
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07068
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch diskette
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: DMA V2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/434, 120
 FILING DATE: 514
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/297, 950
 FILING DATE: 08/22/93
 APPLICATION NUMBER: US/08/226, 108
 FILING DATE: 08/22/93
 APPLICATION NUMBER: US/07/937, 462
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Olslein, Elliot M.
 REGISTRATION NUMBER: 24, 025
 REFERENCE/DOCKET NUMBER: 421250-194

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 114:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 26 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Aphis mellifera
 FEEDBACK: melittin peptide
 NAME/KEY: melittin peptide
 PUBLICATION INFORMATION:
 AUTHORS: Habermann, E.
 AUTHORS: Jentsch, J.
 TITLE: Sequenzanalyse des Melittins aus
 JOURNAL: Hoppe-Seyler's Zeitschrift
 VOLUME: 348
 PAGES: 37-50
 DATE: 1987
 US-08-434-120-114

Query Match 66.0%; Score 93; DB 1; Length 26;
 Best Local Similarity 68.4%; Pred. No. 6.1e-10;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
 Oy 1 GIGAVLXXXXXVLTGTPALISMIXXXXXXXXRQ 38
 Db 1 GIGAVLX---VLTGTPALISMIX---RK---RQ 26

RESULT 13

US-08-465-325-93
 Sequence 93, Application US/08465325
 Patent No. 5686563
 GENERAL INFORMATION:
 APPLICANT: Magalain Pharmaceuticals Inc.
 APPLICANT: 5110 Campus Drive
 TITLE OF INVENTION: Biologically Active Peptides Having
 TITLE OF INVENTION: N-Terminal Substitutions
 NUMBER OF SEQUENCES: 153
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Finegan, Henderson, Parabow, Garrett &
 ADDRESSEE: Dunner
 STREET: 1900 I. Street, N.W. Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465, 325
 FILING DATE: 05-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/184, 462
 FILING DATE: 08/18/94
 APPLICATION NUMBER: 07/891, 201
 FILING DATE: 01-JUN-92
 ATTORNEY/AGENT INFORMATION:
 NAME: Fordis, Jean B
 REGISTRATION NUMBER: 32,984
 REFERENCE/DOCKET NUMBER: 05387, 0021-03000

```

Matches      26:  Conservative      0:  Mismatches      0:  Indels      12:  Caps      3
OY      1 GIGAVLXXXXXVLTGTGALPILSMIKXXXXRRXXXXR00 38
      |||||||
Db      1 GIGAVLK-----VLTGTGALPILSMIK-----RK-----R00 26

RESULT 15
US-08-039-557-5
Sequence 5, Application US/08039557
Patent No. 5714467
GENERAL INFORMATION:
APPLICANT: Boman, Hans G.
INVENTOR: Boman, Hans G.; Boman, David B.
APPLICANT: Andrew, David
TITLE OF INVENTION: ANTIBACTERIAL AND ANTIMALARIAL HYBRID
TITLE OF INVENTION: PEPTIDES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
ATTORNEY: Jackson, Esq.
STREET: Hackensack Ave., Continental Plaza, 4th
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY disk
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,557
FILING DATE: 26-MAR-1993
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
PRIORITY APPLICATION NUMBER: US 07/831,462
FILING DATE: 05-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/449,593
FILING DATE: 12-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,777
FILING DATE: 05-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FROM OTHER SOURCES:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
BIOPHYSICAL: YES
HYDROPHETICAL: YES
POSTSCRIPT TYPE: N-terminal
US-08-039-557-5
Query Match
Best Local Similarity 68.0%; Score 93; DB 1; Length 26;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Caps 3
OY      1 GIGAVLXXXXXVLTGTGALPILSMIKXXXXRRXXXXR00 38
      |||||||
Db      1 GIGAVLK-----VLTGTGALPILSMIK-----RK-----R00 26

Search completed: May 9, 2003, 16:19:27
Job time : 14.7843 secs

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:15:46 ; Search time 16,0196 Seconds
(without alignments)
218,294 Million cell updates/sec

Title: US-09-851-422a-2
Perfect score: 141
Sequence: 1 GIGAVLKKXXXVLTGIPALISWIKXXXXXXXRQ 38

Scoring table: BIOSUM62
Gapext 10.0 ; Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications-AA*

1: /cgn2.6/p/odata/2/pubpaa/US08_NEW_PUB_PEP.*
2: /cgn2.6/p/odata/2/pubpaa/PCT_NEW_PUB_PEP.*
3: /cgn2.6/p/odata/2/pubpaa/US06_NEW_PUB_PEP.*
4: /cgn2.6/p/odata/2/pubpaa/US06_PUBCOM_PEP.*
5: /cgn2.6/p/odata/2/pubpaa/US06_NEW_PUB_PEP.*
6: /cgn2.6/p/odata/2/pubpaa/PCTUS_PUBCOM_PEP.*
7: /cgn2.6/p/odata/2/pubpaa/PCTUS_PUBCOM_PEP.*
8: /cgn2.6/p/odata/2/pubpaa/US08_PUBCOM_PEP.*
9: /cgn2.6/p/odata/2/pubpaa/US09_NEW_PUB_PEP.*
10: /cgn2.6/p/odata/2/pubpaa/US09_PUBCOM_PEP.*
11: /cgn2.6/p/odata/2/pubpaa/US10_NEW_PUB_PEP.*
12: /cgn2.6/p/odata/2/pubpaa/US10_PUBCOM_PEP.*
13: /cgn2.6/p/odata/2/pubpaa/US06_NEW_PUB_PEP.*
14: /cgn2.6/p/odata/2/pubpaa/US06_PUBCOM_PEP.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	66.0	26	9	US-09-999-745-50
2	93	66.0	26	9	US-09-554-000-74
3	93	66.0	26	9	US-09-820-0534-72
4	93	66.0	26	9	US-09-847-208-76
5	92	66.0	26	9	US-09-752-593-1
6	92	65.2	26	10	US-09-917-340-25
7	87.5	62.1	27	10	US-09-030-619-120
8	72	51.1	26	10	US-09-030-619-120
9	64	45.4	28	10	US-09-909-652-7
10	64	45.4	28	10	US-09-030-619-121
11	63	44.7	25	10	US-09-756-2504-3
12	63	44.7	25	10	US-09-030-619-124
13	51	36.2	26	10	US-09-030-619-128
14	51	36.2	26	10	US-09-030-619-130
15	50	35.5	26	10	US-09-820-0534-69
16	48	34.0	23	9	US-09-756-2504-2
17	48	34.0	13	10	US-10-208-948-2
18	47	33.3	1603	9	US-10-208-948-2
19	47	33.3	2159	9	US-10-208-948-2

20	45	31.9	26	10	US-09-030-619-125	Sequence 125, App
21	43	30.5	26	10	US-09-030-619-123	Sequence 123, App
22	43	30.5	425	10	US-09-798-029-2	Sequence 2, App1
23	42	29.8	110	0	US-09-798-029-18	Sequence 18, App1
24	42	29.8	110	0	US-09-798-029-18	Sequence 18, App1
25	41	29.1	275	12	US-10-014-927-16	Sequence 30, App1
26	41	29.1	529	9	US-10-051-902-28	Sequence 28, App1
27	41	29.1	529	9	US-10-051-902-28	Sequence 28, App1
28	40.5	28.7	761	9	US-09-147-947-4	Sequence 4, App1
29	40	28.4	345	10	US-09-943-671-3	Sequence 3, App1
30	40	28.4	4150	10	US-09-808-880-2	Sequence 4, App1
31	39.2	26.0	416	0	US-09-927-6168-4	Sequence 4, App1
32	39.2	26.0	416	0	US-09-927-6168-4	Sequence 4, App1
33	39	27.7	197	10	US-09-815-282-13736	Sequence 13736, App1
34	39	27.7	316	10	US-09-115-150-5	Sequence 5, App1
35	39	27.7	389	9	US-10-081-816-21	Sequence 21, App1
36	39	27.7	539	9	US-10-051-902-26	Sequence 26, App1
37	38.5	27.3	539	9	US-10-051-902-26	Sequence 26, App1
38	38.5	27.3	663	10	US-09-991-630-25	Sequence 25, App1
39	38.5	27.3	2597	10	US-09-505-129-10	Sequence 10, App1
40	38.5	27.3	2597	10	US-09-505-129-10	Sequence 10, App1
41	38.5	27.3	2597	10	US-09-991-630-25	Sequence 25, App1
42	38.5	27.3	2597	10	US-09-991-630-25	Sequence 25, App1
43	38.5	27.3	2597	10	US-09-991-630-25	Sequence 25, App1
44	38.5	27.3	2597	10	US-09-991-630-25	Sequence 25, App1
45	38	27.0	17	10	US-09-859-377-2	Sequence 2, App1

ALIGNMENTS

RESULT 1
US-09-999-745-50
Sequence 50, Application US/0999745
Patent No. US20020157120A1
APPLICANT: TALEN, ROGER Y.
APPLICANT: TALEN, ROGER Y.
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
CURRENT FILING DATE: 2001-10-23
PRIORITY FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50
LENGTH: 26
ORGANISM: Aphis mellifera
US-09-999-745-50
Query Match 66.0% ; Score 93 ; DB 9 ; Length 26 ;
Best Local Similarity 68.4% ; Pred. No. 4.9e-09 ;
Matches 26 ; Conservative 0 ; Mismatches 0 ; Indels 12 ; Gaps 3 ;
DB 1 GIGAVLKKXXXVLTGIPALISWIKXXXXXXXRQ 38
US-09-554-000-74
Sequence 34, Application US/09554000
Patent No. US2002016364A1
APPLICANT: TALEN, ROGER Y.
APPLICANT: TALEN, ROGER Y.
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
FILE REFERENCE: DETECTION OF ANALYTES
CURRENT APPLICATION NUMBER: US/09/554.000

```

: CURRENT FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: 08/918,252
: PRIOR FILING DATE: 1997-03-14
: NUMBER OF SEQ NOS: 30
: SOFTWARE: FASTSD for Windows version 4.0
: SEQ ID NO 34
: LENGTH: 26
: TYPE: PRT
: ORGANISM: Apis mellifera
: US-09-554-000-34

Query Match
: Score 93: DB 9: Length 26:
: Best Local Similarity 68.4%: Pred. No. 4.9e-09:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGAVKXXXXXVLTGIPALISMIXKXXXXXHQ 38
DB 1 GIGAVK----VLTGIPALISMIRK----RQ 26

RESULT 3
: US-09-820-053A-72
: Sequence 72: Application US/09820053A
: Publication No. US200308342A1
: GENERAL INFORMATION:
: APPLICANT: Owen, Donald R
: TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
: FILE REFERENCE: HELX027
: CURRENT APPLICATION NUMBER: US/09/820_053A
: CURRENT FILING DATE: 2001-03-28
: NUMBER OF SEQ ID NOS: 165
: SOFTWARE: Patencin Ver. 2.1
: SEQ ID NO 26
: LENGTH: 26
: TYPE: PRT
: ORGANISM: ARTIFICIAL SEQUENCE
: FEATURE:
: OTHER INFORMATION: SYNTHETIC SEQUENCE
: NAME/KEY: MOD_RES
: LOCATION: (26)
: OTHER INFORMATION: AMIDATION
: US-09-820-053A-72

Query Match
: Score 93: DB 9: Length 26:
: Best Local Similarity 68.4%: Pred. No. 4.9e-09:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGVXXXXXXXXXVLTGIPALISMIXKXXXXXHQ 38
DB 1 GIGVAK-----VLTGIPALISMIRK----RQ 26

RESULT 4
: US-09-030-619-216
: Sequence 216: Application US/090306198
: Publication No. US2003063061A1
: GENERAL INFORMATION:
: APPLICANT: Krieger, Timothy J.
: APPLICANT: Taylor, Robert
: APPLICANT: Erile, Douglas
: APPLICANT: Fraser, Janet R.
: APPLICANT: West, Michael H.P.
: APPLICANT: McNICOLL, Patricia J.
: TITLE OF INVENTION: METHODS OF USING AND METHODS FOR TREATING
: TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
: FILE REFERENCE: 660081.406
: CURRENT APPLICATION NUMBER: US/09/030_619B
: CURRENT FILING DATE: 1998-02-25
: NUMBER OF SEQ ID NOS: 232
: SOFTWARE: FASTSD for Windows version 3.0
: SEQ ID NO 216
: LENGTH: 26
```

```

: TYPE: PRT
: ORGANISM: Apis mellifera
: US-09-030-619-216

Query Match
: Score 93: DB 10: Length 26:
: Best Local Similarity 68.4%: Pred. No. 4.9e-09:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGVXXXXXXXXVLTGIPALISMIXKXXXXXHQ 38
DB 1 GIGAVK----VLTGIPALISMIRK----RQ 26

RESULT 5
: US-09-847-208-26
: Sequence 26: Application US/09847208
: Publication No. US2003082190A1
: GENERAL INFORMATION:
: APPLICANT: Zhou, Andrew
: APPLICANT: Zhou, Andrew
: APPLICANT: Zhou, Andrew
: APPLICANT: Zhou, Andrew
: APPLICANT: Zhou, Andrew
: TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
: TITLE OF INVENTION: 19E-MEDIATED ALLERGIC DISEASES
: FILE REFERENCE: UC67_002A
: CURRENT APPLICATION NUMBER: US/09/847_208
: CURRENT FILING DATE: 2001-05-01
: NUMBER OF SEQ ID NOS: 1
: SOFTWARE: FastSD for Windows version 4.0
: SEQ ID NO 26
: LENGTH: 70
: TYPE: PRT
: ORGANISM: Apis mellifera(Honeybee)Apis cerana(Ind. Honeybee)
: US-09-847-208-26

Query Match
: Score 93: DB 9: Length 70:
: Best Local Similarity 68.4%: Pred. No. 1.4e-08:
Matches 26: Conservative 0: Mismatches 0: Indels 12: Gaps 3:

OY 1 GIGVXXXXXXXXVLTGIPALISMIXKXXXXXHQ 38
DB 44 GIGAVK----VLTGIPALISMIRK----RQ 69

RESULT 6
: US-09-752-533-1
: Sequence 1, Application US/09752533
: Patent No. US20020107389A1
: GENERAL INFORMATION:
: APPLICANT: Livingston, Stephen M.
: APPLICANT: Jones, David S.
: APPLICANT: Livingston, Douglas A.
: APPLICANT: Yu, Lin
: TITLE OF INVENTION: CONJUGATES OF CHEMICALLY DEFINED
: TITLE OF INVENTION: NON-POLYMERIC VALENCY PLATFORM MOLECULES AND BIOLOGICALLY
: TITLE OF INVENTION: ACTIVE MOLECULES
: CURRENT APPLICATION NUMBER: US/09/752_533
: CURRENT FILING DATE: 2000-12-25
: PRIOR APPLICATION NUMBER: 08/769,041
: PRIOR FILING DATE: 1996-12-18
: PRIOR APPLICATION NUMBER: 08/453,254
: PRIOR FILING DATE: 1995-05-30
: PRIOR APPLICATION NUMBER: 08/452,506
: PRIOR FILING DATE: 1995-05-30
: PRIOR APPLICATION NUMBER: 07/914,869
: PRIOR FILING DATE: 1992-07-15
: PRIOR APPLICATION NUMBER: 08/118,055
: PRIOR FILING DATE: 1993-09-08
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSD for Windows version 4.0
: SEQ ID NO 1
: LENGTH: 26
: TYPE: PRT
```

ORGANISM: Apis cerana
FEATURE:
NAME/RES: AMIDATION
SEQUENCE: 6
OTHER INFORMATION: Gln attached to an Amide
US-09-752-533-1

Query Match
Best Local Similarity 65.2%; Score 92; DB 10; Length 26;
Best Local Similarity 84.0%; Pred. No. 7.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 1 GIGAVLXXXXVITGTPALISWK 25
DB 1 GIGAVLK---VITGTPALISWK 21

RESULT 7
US-09-917-340-25
Sequence 25; Application US/09917340
Patent No. US20020090369A1
APPLICANT: Murphy, Christopher J.
APPLICANT: McNulty, Jonathan F.
APPLICANT: Reid, Ted W.
TITLE OF INVENTION: Transplant Media
FILE REFERENCE: PTLANT-60468
CURRENT APPLICATION NUMBER: US/09/917,340
PRIORITY FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/221,632
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/249,602
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/290,932
NUMBER OF SEQ ID NOS: 96
SOFTWARE: FASTSEQ for Windows Ver. 2.0
SEQ ID NO 25
LENGTH: 27
TYPE: PRP
ORGANISM: Apis mellifera
US-09-917-340-25

Query Match
Best Local Similarity 62.1%; Score 87.5; DB 10; Length 27;
Best Local Similarity 65.8%; Pred. No. 3e-09;
Matches 25; Conservative 0; Mismatches 2; Indels 11; Gaps 3;

QY 1 GIGAVLXXXXVITGTPALISWKXXXXXRRQ 38
DB 1 GIGAVLK---VITGTPALISWK---SRK---KROQ 27

RESULT 8
US-09-030-619-120
Sequence 120; Application US/090306198
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Fraser, Janet R.
APPLICANT: Praser, Janet R.
APPLICANT: McNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTI-BIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,6198
PRIORITY FILING DATE: 1998-02-25
CURRENT FILING DATE: 2000-07-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 120
LENGTH: 26
TYPE: PRP
ORGANISM: Artificial Sequence

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-120

Query Match
Best Local Similarity 51.1%; Score 72; DB 10; Length 26;
Best Local Similarity 81.8%; Pred. No. 1.7e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 1 GIGAVLXXXXVITGTPALIS 22
DB 9 GIGAVLK---VITGTPALIS 26

RESULT 9
US-09-909-652-7
Sequence 7; Application US/09909652
Patent No. US20020025537A1
GENERAL INFORMATION:
APPLICANT: Ratios Scientific, Inc.
APPLICANT: Coleman, William J.
APPLICANT: Youvan, Douglas C.
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR GENERATING
TITLE OF INVENTION: AND SCREENING COMPOUNDS THAT AFFECT CELL VIABILITY
FILE REFERENCE: 22346-7001
CURRENT APPLICATION NUMBER: US/09/909,652
PRIORITY FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 60/419,179
PRIOR FILING DATE: 2000-07-19
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7
LENGTH: 28
TYPE: PRP
FEATURE: Artificial Sequence
OTHER INFORMATION: CEWA analogue of cecropin-melittin hybrid
US-09-909-652-7

Query Match
Best Local Similarity 45.4%; Score 64; DB 10; Length 28;
Best Local Similarity 80.0%; Pred. No. 0.00042;
Matches 16; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 GIGAVLXXXXVITGTPAL 20
DB 9 GIGAVLK---VITGTPAL 24

RESULT 10
US-09-030-619-121
Sequence 121; Application US/090306198
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Fraser, Janet R.
APPLICANT: Praser, Janet R.
APPLICANT: McNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTI-BIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,6198
PRIORITY FILING DATE: 1998-02-25
CURRENT FILING DATE: 2000-07-25
NUMBER OF SEQ ID NOS: 121
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 121
LENGTH: 28
TYPE: PRP
ORGANISM: Artificial Sequence

OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-121

Query Match 45.4%; Score 64; DB 10; Length 28;
Best Local Similarity 80.0%; Pred. No. 0.00042;
Matches 16; Conservative 0; Mismatches 0; Indels 4; Caps 1;

QY 1 GIGAVLXXXXVLTGRLPAL 20
DB 9 GIGAVLX----VLTGRLPAL 24

RESULT 11

US-09-756-2508-3
Sequence 3, Application US/097562508
Patent No. US2005058020A1

GENERAL INFORMATION: GIGAVLXXXXXVLTGRLPAL 20
TITLE OF INVENTION: Isolated amphiphilic peptides derived from the cytoplasmic tail of
FILE REFERENCE: 4-306634/CI/CTI 2012
CURRENT APPLICATION NUMBER: US/09/756,2508
CURRENT FILING DATE: 2001-12-13
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 3.0
LENGTH: 15
SEQ ID NO: 122
TYPE: PART

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Shortened analogue of melittin peptide

NAME/KEY: MUTAGEN

LOCATION: (1)...(15)

US-09-730-2508-3

Query Match 44.7%; Score 63; DB 10; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.00031;
Matches 12; Conservative 1; Mismatches 0; Indels 0;

QY 12 VLTGRLPALISM 24
DB 3 VLTGRLPALISM 15

RESULT 12

US-09-030-619-122
Sequence 12, Application US/090306198
Patent No. US20020035061A1

GENERAL INFORMATION: VLTGRLPALISM 24
TITLE OF INVENTION: Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: MCNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,6198
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 122
TYPE: PART

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Cationic Peptide Analogue

US-09-030-619-122

Query Match 36.2%; Score 51; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.06;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 12 VLTGRLPALIS 22
DB 16 VLTGRLPALIS 26

RESULT 13

US-09-030-619-124
Sequence 124, Application US/090306198
Patent No. US20020035061A1

GENERAL INFORMATION: VLTGRLPALIS 22
TITLE OF INVENTION: Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: MCNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,6198
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
LENGTH: 26
SEQ ID NO: 124
TYPE: PART

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Cationic Peptide Analogue

US-09-030-619-124

Query Match 36.2%; Score 51; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 11; Conservative 0; Mismatches 0; Indels 0;

QY 12 VLTGRLPALIS 22
DB 16 VLTGRLPALIS 26

RESULT 14

US-09-030-619-128
Sequence 128, Application US/090306198
Patent No. US20020035061A1

GENERAL INFORMATION: VLTGRLPALIS 26
TITLE OF INVENTION: Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: MCNICOL, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,6198
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 128
TYPE: PART

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Fusion Peptides

LOCATION: (4)...(4)

OTHER INFORMATION: Xaa = Hydrophilic Residue

NAME/KEY: VARIANT

LOCATION: (5)...(6)

```

OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (7)...(8)
OTHER INFORMATION: xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (9)...(9)
OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (10)...(11)
OTHER INFORMATION: xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (12)...(13)
OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (14)...(15)
OTHER INFORMATION: xaa - Hydrophilic Residue
US-09-030-619-128

Query Match
US-09-030-619-130 36.2% Score 51: DB 10: Length 26:
Best Local Similarity 100.0%: Pred.No. 0.06:
Matches 11: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 12 VLTGLPALIS 22
DB 16 VLTGLPALIS 26

```

```

RESULT 15
US-09-030-619-130
Sequence 130: Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Kriesger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Taylor, Robert
APPLICANT: Taylor, Robert
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McInicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
TITLE OF INVENTION: WITH ANTIBIOTICS
FILE NO. OF INVENTION: 660081-06 US/09/030,619B
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 130
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Fusion Peptides
NAME/KEY: VARIANT
LOCATION: (4)...(4)
OTHER INFORMATION: xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (5)...(6)
OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (7)...(8)
OTHER INFORMATION: xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (9)...(9)
OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (10)...(11)
OTHER INFORMATION: xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (12)...(13)
OTHER INFORMATION: xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (14)...(15)
OTHER INFORMATION: xaa - Hydrophilic Residue

```

```

US-09-030-619-130
Query Match
US-09-030-619-130 35.5% Score 50: DB 10: Length 26:
Best Local Similarity 90.9%: Pred.No. 0.089:
Matches 10: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 12 VLTGLPALIS 22
DB 16 VLTGLPALIS 26

Search completed: May 9, 2003, 16:20:17
Job time : 17.0196 secs

```

THIS PAGE BLANK (USPTO)

29	42.5	30.1
----	------	------

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	93	66.0	26	1	MEMBR1	melittin - melittin
2	93	66.0	70	1	MPBR1	melittin - melittin
3	87.5	62.1	27	1	MEMB2	melittin - melittin
4	87	61.7	26	1	MEMB3	melittin - melittin
5	82	58.2	26	1	MEMBFC	melittin - melittin
6	46	34.6	64.2	1	146534	renal sodium-dep
7	45	33.9	37	2	AT0115	conserved hypot
8	45	31.9	37	2	AT0115	class I histocomp
9	45	31.9	35.7	2	AT0115	class I histocomp
10	45	31.9	35.7	2	AT0115	class I histocomp
11	45	31.9	35.7	2	AT0115	class I histocomp
12	45	31.9	35.7	2	AT0119	class I histocomp
13	45	31.9	35.5	2	172218	class I histocomp
14	44	31.2	35.2	2	AT0113	class I histocomp
15	44	31.2	35.2	2	AT0113	class I histocomp
16	44	31.2	71.4	2	125056	class I histocomp
17	44	31.2	75.2	2	AT2550	conserved hypot
18	44	31.2	74.6	2	H87466	beta-D-glucosidase
19	44	31.2	43.4	1	A53489	dyerlen heavy chain
20	43.5	30.9	200	2	G75550	conserved hypot
21	43	30.5	174	2	G93083	conserved domain p
22	43	30.5	243	2	AT0201	hypothetical class
23	43	30.5	243	2	AT0201	hypothetical class
24	43	30.5	303	2	F68156	probable amino ac
25	43	30.5	463.2	2	E91156	probable amino ac
26	43	30.5	463.2	2	B86002	probable amino ac
27	43	30.5	462	2	F65131	hypothetical 47.5
28	43	30.5	744	2	CG7036	transport protein
29	42.5	30.1	330	2	AT81382	probable conserved

RESULT 1
MEMBCI
mellicci, major indian honeybee
C:Date: 13-Jul-1981 ;text_change 15-Oct-1996
C:Date: 13-Jul-1981 ;sequence_revision 13-Jul-1981
C:Accession: A01762
R:Kreil, G.
EESB Lett., 33, 241-244, 1973
A:Title: Structure of mellictin isolated from two species of honey bees.

[illegible]

[illegible]

```

RESULT 3
MEMB2
mellitin: minor - honeybee
C:Species: Apis mellifera (honeybee)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
Accession: A01764

```

R. Habermann, E. Jentsch. *J. Physiol. Chem.* **348**, 37-50, 1967
Höppe-Seyler's Z. *Biochem. Anal.* **15**, 1-3, 1904
A. Fülle: Sequenzanalyse des Melittins aus den Deutschen Spaltstücken
M. Kossuth, R. W. Bland. *Mol. Cell Biochem.* **101**/764 A194d0; PMID:65327913; PMID:5592400
A.: Peptide analysis. *Methods Enzymol.* **101**/764
A.: Residues: 1-27 <FAB>
R. Schröder, E.: Luebeck, K.; Lehmann, M.; Boettl, I.
Experientia **27**, 764-765, 1971
A. Fülle: Hemolytic activity and action on the surface tension of aqueous solutions O
A. Contents: annotation: synthesis
C: Superfamily: melittin major
C: Keywords: amidated carboxyl end; hemolysis; (dim) tetraetars venom
Z/27/Modified site: amidated carboxyl end; (dim) tetraetars experimental

```

RESULT 4
METHOD CD
Query Match      - giant honeybee (tentative sequence)
Best Local Similarity   76.0%; Score 87; DB 1; Length 26;
Matches          19; Conservative    2; Mismatches    0; Indels    4; Gaps    1;
QY              1 CIGATGXXXXXXXXXXVLTGTCTGALTSWIK 25
Db              1 CIGATLK-----VLTSTGTCTGALTSWIK 21

```

```

Result 5
MEHBCF
    mellittin - little honeybee
C:species: Apis mellifera Floren (little honeybee)
C:species: Jun-Jun-1981
C:accession: A01765
R:protein: G
A:Title: Structure of mellittin isolated from two species of honey bees.
A:reference number: A01392
A:accession: A01765
A:protein: mellittin
A:residues: 126
C:superfamily: mellittin major
C:keywords: amidated carboxyl end; hemolytic; homotetramer; venom
2:56/Modified site: amidated carboxyl end; (Gln) status experimental

```

Query Match	58.28;	Score 82;	DB 1;	Length 26;
Best Local Similarity	72.0%;	Pred. No. 4	8e-07;	
Matches 18;	Conservative 1;	Mismatches 2;	Indels 4;	Gaps 1


```

OY 1 GIGAVLKKXXXVLTGCLPALISMIR 25
    1111:11 11 1111 111111
DB 1 GIGALIK-----VLATGSLPLLSMWK 21

RESULT 6
146534
renal sodium-dependent phosphate transporter type II - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: S11137
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.;
Am. J. Physiol. 266, F623-F633, 1995
A:Title: Cloning of a rabbit renal Na-Pi cotransporter, which is regulated by dietary ph
A:Reference number: 146534; PMID:95251033; PMID:773319
A:Accession: S11534
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Cross-references: EMBL:U020793; NID:9687736; PID:NAA7682.1; PID:9687737
Query Match 32.6%; Score 46; DB 2; Length 642;
Best Local Similarity 40.9%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 3 GAVLKKXXXVLTGCLPALISMIR 24
    1111:11 11 1111 111111
DB 382 GGVAKYONVITNDLPARTNV 403

RESULT 7
A10113
conserved hypothetical protein YPO0945 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 25-Feb-1994 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: A10113
R:Perkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Penrice, M.B.;
demo-Tarrago, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: 11585360; PMID:11585360
A:Accession: A10113
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <KUR>
A:Cross-references: GB:AL50842; PID:CA89788.1; PID:915979015; GSPED:GN0175
C:Genetics:
A:Gene: YPO0945
C:Superfamily: Mechanococcus jansaschii conserved hypothetical protein M00226
Query Match 31.9%; Score 45; DB 2; Length 197;
Best Local Similarity 52.9%; Pred. No. 5;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 AVLKXXXVLTGCLPAL 20
    1111:11 11 1111 111111
DB 50 AILKAKHAAHMTGIDRAL 66

RESULT 8
S11137
class I histocompatibility antigen alpha chain - cotton-top tamarin
C:Species: Saguinus oedipus (cotton-top tamarin)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
C:Accession: S11137
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.;
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hominoid
A:Reference number: S10934; PMID:90309971; PMID:2114550
C:Accession: S11137
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA

```

```

A:Residues: 1-357 <MAT>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F:212-277/Domain: immunoglobulin homology <IM>
Query Match 31.9%; Score 45; DB 2; Length 357;
Best Local Similarity 29.7%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKKXXXVLTGCLPALISMIRKXXXXX 37
    1111:11 11 11 11 11 11
DB 302 GVALALILGAVTGAVVAAMVMMKRSDDKGGSYSD 338

RESULT 9
S11134
class I histocompatibility antigen alpha chain - cotton-top tamarin
C:Species: Saguinus oedipus (cotton-top tamarin)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
C:Accession: S11134
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.;
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hominoid
A:Reference number: S10934; PMID:90309971; PMID:2114550
C:Accession: S11134
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-357 <MAT>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F:212-277/Domain: immunoglobulin homology <IM>
Query Match 31.9%; Score 45; DB 2; Length 357;
Best Local Similarity 29.7%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKKXXXVLTGCLPALISMIRKXXXXX 37
    1111:11 11 11 11 11 11
DB 302 GVALALILGAVTGAVVAAMVMMKRSDDKGGSYSD 338

```

```

RESULT 10
S11135
class I histocompatibility antigen alpha chain - cotton-top tamarin
C:Species: Saguinus oedipus (cotton-top tamarin)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
C:Accession: S11135
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.;
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hominoid
A:Reference number: S10934; PMID:90309971; PMID:2114550
C:Accession: S11135
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-357 <MAT>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F:212-277/Domain: immunoglobulin homology <IM>
Query Match 31.9%; Score 45; DB 2; Length 357;
Best Local Similarity 29.7%; Pred. No. 11;
Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKKXXXVLTGCLPALISMIRKXXXXX 37
    1111:11 11 11 11 11 11
DB 302 GVALALILGAVTGAVVAAMVMMKRSDDKGGSYSD 338

RESULT 11
S11140
class I histocompatibility antigen alpha chain - cotton-top tamarin
C:Species: Saguinus oedipus (cotton-top tamarin)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
C:Accession: S11140
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.;
Nature 346, 60-63, 1990

```

A>Title: Evolution of the MHC class I genes of a New World primate from ancestral homoid
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S11139
 A:Species: *Saguinus oedipus* (cotton-top tamarin)
 C:Dates: 25-Feb-1994 #sequence, revision 26-May-1995 #text, change 23-Jul-1999
 C:Accession: S11133
 R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homoid
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S11139
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trn
 A:Molecule type: mRNA
 A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: Immunoglobulin homology <IMX>

Query Match 31.9%; Score 45; DB 2; Length 357;
 Best local similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

DB 302 GIVALLALIGVVTGAVVAAMVWKKSSDKKCGSYSO 338

RESULT 12
 S11139
 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: *Saguinus oedipus* (cotton-top tamarin)
 C:Dates: 25-Feb-1994 #sequence, revision 26-May-1995 #text, change 23-Jul-1999
 C:Accession: S11133
 R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homoid
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S11139
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trn
 A:Molecule type: mRNA
 A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: Immunoglobulin homology <IMX>

Query Match 31.9%; Score 45; DB 2; Length 357;
 Best local similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

DB 302 GIVALLALIGVVTGAVVAAMVWKKSSDKKCGSYSO 338

RESULT 13
 172218
 class I histocompatibility antigen I-C alpha chain precursor - cotton-top tamarin
 N:Alternate names: MHC class IC protein
 C:Species: *Saguinus oedipus* (cotton-top tamarin)
 C:Dates: 02-Jul-1999 #sequence, revision 02-Jul-1996 #text, change 23-Jul-1999
 C:Accession: 172218; S11142
 R:Watkins, D.I.; Letvin, N.L.; Hughes, A.L.; Tedder, T.F.
 J:Immunol. 144, 1136-1143, 1990
 A:Title: Molecular cloning of cDNA that encode MHC class I molecules from a new world pr
 A:Reference number: 156053; WUID:90111120; PMID:2104912
 A:Accession: 172218
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-365 <RESS>
 A:Cross-references: GB:M3477; NID:9343319; PIDN:AAA6955.1; PID:9343320
 R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
 Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homoid
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S10934
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual trn
 A:Molecule type: mRNA
 A:Residues: 9-365 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: glycoprotein; transmembrane protein
 F:220-285/Domain: Immunoglobulin homology <IMX>

Query Match 31.9%; Score 45; DB 2; Length 365;

Best Local Similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

DB 310 GIVALLALIGVVTGAVVAAMVWKKSSDKKCGSYSO 346

RESULT 14
 S11133
 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: *Saguinus oedipus* (cotton-top tamarin)
 C:Dates: 25-Feb-1994 #sequence, revision 26-May-1995 #text, change 23-Jul-1999
 C:Accession: S11133
 R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
 Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S11133
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: Immunoglobulin homology <IMX>

Query Match 31.2%; Score 44; DB 2; Length 357;
 Best local similarity 29.7%; Pred. No. 15;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

DB 302 GIVALLALIGVVTGAVVAAMVWKKSSDKKCGSYSO 338

RESULT 15
 156053
 class I histocompatibility antigen I-A alpha chain precursor - cotton-top tamarin
 N:Alternate names: MHC class IA protein
 C:Species: *Saguinus oedipus* (cotton-top tamarin)
 C:Dates: 02-Jul-1999 #sequence, revision 02-Jul-1996 #text, change 23-Jul-1999
 C:Accession: 156053
 R:Watkins, D.I.; Letvin, N.L.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.
 J:Immunol. 144, 1136-1143, 1990
 A:Title: Molecular cloning of cDNA that encode MHC class I molecules from a new world
 A:Reference number: 156053; WUID:90111120; PMID:2104912
 A:Accession: 156053
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-365 <RESS>
 A:Cross-references: GB:M3475; NID:9343315; PIDN:AAA6955.1; PID:9343316
 R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
 Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
 A:Reference number: S10934; WUID:90309971; PMID:2114550
 A:Accession: S10934
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 9-365 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 C:Keywords: glycoprotein; transmembrane protein
 F:220-285/Domain: Immunoglobulin homology <IMX>

Query Match 31.2%; Score 44; DB 2; Length 365;
 Best local similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

DB 310 GIVALLALIGVVTGAVVAAMVWKKSSDKKCGSYSO 346

Search completed: May 9, 2003, 16:18:46
 Job time: 25.8431 secs

THIS PAGE BLANK (USPTO)

RA "The structure of melittin, 11. Interpretation of the structure.";
 RL J Biol. Chem. 257:6016-6022(1982).
 RM STRUCTURE BY NMR OF 44-69.
 RC SPECIES-A: mellifera;
 RA Barnham K.J., Hewish D., Wermkester J., Curran C., Kirkpatrick A.,
 RA Barton N., Norton R., Rivett D.;
 RA Submitted (JUN-1998) to the PDB data bank.
 RM 181791.
 RX MEDLINE=90254148; PubMed=2187556.
 RA Dempsey C.E.;
 RL "The actions of melittin on membranes.";
 RL Biochim. Biophys. Acta 1031:143-161(1990).
 CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
 CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
 CC PHOSPHOLIPIDS. IT INHIBITS WELL-KNOWN TRANSPORT PUMPS SUCH AS THE
 CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
 CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
 CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.
 CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER.
 CC -1- CELLULAR LOCATION: SECRETED.
 CC -1- DATABASE: N-TERMINAL-MELITTIN MAJOR HAS 80% OF THE ACTIVITY OF
 CC -1- NOTE-Issue 12 of July 2001:
 CC WWW="http://www.expasy.org/spolight/articles/spclt012.html".
 CC -1- THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC by non-profit institutions as long as its contents are made
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 DR EMBL: A02007; CAA26038.1;
 DR EMBL: A02007; CAA26038.1;
 DR PIR: A01764; MEHBC1.
 DR PIR: A01762; MEHBC1.
 DR PDB: 2MLT; 15-JUL-92.
 DR PDB: 1BML; 06-JAN-99.
 DR InterPro: IPR002116; Melittin.
 DR Pfam: PF01372; Melittin; 1.
 DR ProDom: PD014636; Melittin; 1.
 KW Aminoacids; Toxin; Amidation; Formylation; Signal; 3D-structure;
 KW Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 43
 FT CHAIN 44 69
 FT MOD_RES 44 44
 FT MOD_RES 69 69
 FT VARIANT 64 64
 FT VARIANT 67 70
 FT HELIX 45 53
 FT TURN 54 54
 FT HELIX 55 68
 SO SEQUENCE 70 AA: 7585 MW: 60755CO91C33BB6 CRC64;
 Query Match 66.0% Score 93; DB 1; Length 70;
 Best local similarity 68.4% Pseq No. 6.4e-09;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
 DB 1 GIGAVLXXXXXVLTGCPALISWIKXXXXXKXXXXXKQ 38
 DB 44 GIGAVLK-----VLTGCPALISWIK-----RK-----RQ 65
 RESULT 2
 MEL_APIDO

ID MEL_APIDO STANDARD; PRT; 26 AA.
 AC E01502; 1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUN-2002 (Rel. 41, Last annotation update)
 DE Melittin.
 OS Apis dorsata (Giant honeybee).
 CC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 CC Hymenoptera; Apidae; Apidae; Apis.
 RX MEDLINE=75183194; PubMed=1093875;
 RA Krell G.;
 RA "The structure of Apis dorsata melittin: phylogenetic relationships
 RA between honeybees and dedicated from sequence data.";
 RL J Biol. Chem. 261:1191-1196(1986).
 CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
 CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
 CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
 CC PHOSPHOLIPIDS. IT INHIBITS WELL-KNOWN TRANSPORT PUMPS SUCH AS THE
 CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
 CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
 CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.
 CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER (BY SIMILARITY).
 CC -1- CELLULAR LOCATION: SECRETED.
 CC -1- DATABASE: NAME-Protein SpotLight;
 CC -1- NOTE-Issue 12 of July 2001:
 CC WWW="http://www.expasy.org/spolight/articles/spclt012.html".
 DR PIR: A01763; MEHBC1.
 DR EMBL: A02007; CAA26038.1;
 DR EMBL: A02007; CAA26038.1;
 DR PIR: A01764; MEHBC1.
 DR PIR: A01762; MEHBC1.
 DR PDB: 2MLT; 15-JUL-92.
 DR PDB: 1BML; 06-JAN-99.
 DR InterPro: IPR002116; Melittin.
 DR Pfam: PF01372; Melittin; 1.
 DR ProDom: PD014636; Melittin; 1.
 KW Hemolysis; Toxin; Amidation.
 FT MOD_RES 26
 FT SEQUENCE 26 AA: 2848 MW: 61486F95241E01C CRC64;
 Query Match 61.7% Score 87; DB 1; Length 26;
 Best local similarity 76.0% Pseq No. 2.5e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 DB 1 GIGAVLXXXXXVLTGCPALISWIK 25
 DB I GIGAVLK-----VLTGCPALISWIK 21
 RESULT 3
 MEL_APFL STANDARD; PRT; 26 AA.
 ID MEL_APFL
 AC A01504;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 21-JUN-2002 (Rel. 41, Last annotation update)
 DE Melittin.
 OS Apis florea (Little honeybee).
 CC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 CC Hymenoptera; Apidae; Apidae; Apis.
 RX MEDLINE=7463;
 RA NCB1_TaxID=7463;
 RA (1)
 RA SEQUENCE.
 RA "Structure of melittin isolated from two species of honey bees.";
 RL FEES Lett. 33:241-244(1973).
 CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
 CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
 CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
 CC PHOSPHOLIPIDS. IT INHIBITS WELL-KNOWN TRANSPORT PUMPS SUCH AS THE
 CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
 CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
 CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.

```

CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC DOID: D000012000
CC NAME: NAME-Protein Spotlight.
CC URL: http://www.expasy.org/spotlight/articles/april012.html*.
CC
DR HSP: P01501; MBHL.
DR InterPro: IPR002116; Melittin.
DR Pfam: PF01372; Melittin; 1.
DR ProDom: PD016536; Melittin; 1.
DR Homolysis: Toxin; Amidation.
DR MODRES: 26 AA; 2819 MW; F20E8R82400EF0IC CRC64.
DR SEQUENCE
SQ
Query Match 58.2%; Score 82; DB 1; Length 26;
Best Local Similarity 72.0%; Pred. No. 1.7e-07;
Matches 1; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

OY 1 GCGVYKXXXXVTGGLPILSMIK 25
DB 1 GIGMILK----VLTGTPILSMIK 21

RESULT 4
MLP_RANTE STANDARD: PRT; 22 AA.
ID MLP_RANTE STANDARD: PRT; 22 AA.
CC 556924; 2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Melittin-like peptide (MLP).
OS Rana temporaria (European common frog).
DE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AM Amphibia; Batrachia; Anura; Neobatrachia; Ranidae; Rana.
CC [1] TaxID=9407;
RN [1]
RP SEQUENCE.
RA TISSUE-SKIN;
RX MEDLINE=97175050; PubMed=9022710;
RC Stimaco M., Mignogna G., Canolet S., Miele R., Mangoni M.L.,
RA Bacra D.;
FT Temporal expression profile.
RL Eur J. Biochem. 242:788-792(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: HIGH, TO A.MILITERRA MELITLIN.
KW Amphibian skin; Amidation.
FT MODRES 22 AA; 2313 MW; 0950320E90709B3 CRC64.
SQ
Query Match 36.2%; Score 51; DB 1; Length 22;
Best Local Similarity 45.8%; Pred. No. 0.025;
Matches 1; Conservative 5; Mismatches 4; Gaps 1;

OY 2 ICNAVLEKXXXXVLTGGLPILSMIK 25
DB 2 IGSMILK----VLTGGLPILSMIK 21

RESULT 5
NPT2_RABIT STANDARD: PRT; 642 AA.
ID NPT2_RABIT STANDARD: PRT; 642 AA.
CC 028620;
DT 01-NOV-1997 (Rel. 35, Created)
DT 11-NOV-2002 (Rel. 35, Last sequence update)
DT 11-NOV-2002 (Rel. 35, Last sequence update)
DE Renal sodium-dependent phosphate transporter 2 (Sodium/phosphate
DE cotransporter 2) (Na(+)/Pi cotransporter 2) (Renal sodium-phosphate
DE cotransporter 2) (Renal Na(+)-dependent phosphate cotransporter 2)
DE (NPT-6).
DE SLC34A1 OR SLC17A2.
DR Oryctolagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
CC MEDLINE=95251013; PubMed=7733319;
RX Verri T., Markovitch D., Perego C., North F., Stange G., Sorribas V.,
RA Biber J., Murer H.;
FT Cloning of a rabbit renal Na-Pi cotransporter, which is regulated by
FT dietary phosphate.
FT Am. J. physiol. 268:F626-F633(1995).
CC -1- FUNCTION: NPT 2B IS INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE
CC FROM THE EXTRACELLULAR SPACE INTO THE CELL VIA Na+ CO-TRANSPORT IN THE RENAL BRUSH BORDER
CC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not used for
CC commercial purposes and the source is acknowledged. For commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: U20793; AAA7682.1;
DR InterPro: IPR003841; Na_Pi_cotrans.
DR Pfam: PF02690; Na_Pi_cotrans; 1.
CC TIGR: TIGR01013; 2849.1;
KW Phosphorylation; sodium transport; Transmembrane; Glycoprotein;
CC
FT DOMAIN 1 106
FT TRANSMEM 107 128
FT DOMAIN 129 148
FT TRANSMEM 149 166
FT DOMAIN 167 188
FT TRANSMEM 189 350
FT DOMAIN 351 373
FT TRANSMEM 374 415
FT DOMAIN 416 439
FT TRANSMEM 440 469
FT TRANSMEM 470 490
FT TRANSMEM 491 512
FT TRANSMEM 513 537
FT DOMAIN 538 542
FT TRANSMEM 543 564
FT TRANSMEM 565 642
FT DOMAIN 643 672
FT MODRES 673 711
FT CARBOHYD 712 731
FT CARBOHYD 732 733
FT CARBOHYD 734 735
SQ
Query Match 32.6%; Score 46; DB 1; Length 642;
Best Local Similarity 40.9%; Pred. No. 4.9;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 3 GAVLYKXXXXVLTGGLPILSMIK 24
DB 382 GGVAKVYGVINTDPAPIVW 403

RESULT 6
ICOL_SAOOE STANDARD: PRT; 365 AA.
ID ICOL_SAOOE STANDARD: PRT; 365 AA.
CC P30571; SAOGE
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE Class I histocompatibility antigen, C alpha chain precursor.
DE Saginatus oedipus (Cotton-top tamarin).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
CX NCBI_TaxID:9690.
RN [1]
RX SOURCE FROM N.A.
RY RESOLUTION: 21049120.
RA Watkins D.I., Leavitt N.L., Hughes A.L., Tedder T.F.;
RT "Molecular cloning of cDNA that encodes MHC class I molecules from a
RW New World primate (Saguinus oedipus). Natural selection acts at
RL J. Immunol. 144:1136-1143(1990).
CC CC - FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC CD8+ T CELLS AND IN THE REGULATION OF ALPHA CHAIN AND BETA 2-
CC MICROGLOBULIN.
CC --
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@lsdb.slb.ch see http://www.lsdb.slb.ch/announce/.
CC CC
DR EMBL: M33477; AAA69393.1; .
DR HSSP: Q95352; IHK.
DR InterPro: IPRO03006; IG_MHC.
DR InterPro: IPRO03597; IG_C1.
DR InterPro: IPRO01039; MHC_L1.
DR Pfam: PF00128; MHC_1; 1.
DR ProDom: PD000050; MHC_1; 1.
DR SMART: SM00407; IGC1.1; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KM MHC_I; Transmembrane; Glycoprotein; signal.
FT SIGNAL.
FT FT 25 365
FT STAIL.
FT FT 25 365
FT FT 25 365
FT FT DOMAIN 25 114
FT FT DOMAIN 115 206
FT FT DOMAIN 207 298
FT FT DOMAIN 299 308
FT FT TRANSMEM 309 332
FT FT DOMAIN 333 365
FT FT TRANSMEM 366 383
FT FT DISULFID 227 283
FT FT CARBOHYD 110 110
SQ SEQUENCE 365 AA: 40829 MW: D5A0E97410BA1BE CRC64:
Query Match Best Local Similarity 31.9% Score 45; DB 1; Length 365;
Matches 11; Conservative 55; Mismatches 21; Indels 0; Gaps 0.
OY 1 GIGVIXXXXXXWTCGTATGISIKXXXXXXRRQ 37
| | | | | | | | | | | | | | | | | |
310 GIAYLLALCAAVTGAAVGAAYVAAMVRKKSKDKKCSYSQ 346
| | | | | | | | | | | | | | | | | |
RESULT 7
1 ADI_SNOOE STANDAR: PRF: 365 AA.
1 ADI_SNOOE
AC P30315.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 01-APR-1993 (Rel. 25, Last annotation update)
OS Clasa I histocompatibility antigen, A alpha chain precursor.
DE Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Platyrrhini; Eutelestomi;
OC Neotoma; Cebinae; Primates; Platyrrhini; Callitrichidae; Saguinus.
CX NCBI_TaxID:9690.
RN [1]
RX SOURCE FROM N.A.
RY RESOLUTION: 21049120.
RA Watkins D.I., Leavitt N.L., Hughes A.L., Tedder T.F.;
RT "Molecular cloning of cDNA that encodes MHC class I molecules from a
RW New World primate (Saguinus oedipus). Natural selection acts at
RL J. Immunol. 144:1136-1143(1990).
CC CC - FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC CD8+ T CELLS AND IN THE REGULATION OF ALPHA CHAIN AND BETA 2-
CC MICROGLOBULIN.
CC --
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@lsdb.slb.ch see http://www.lsdb.slb.ch/announce/.
CC CC
DR EMBL: M33477; AAA69393.1; .
DR HSSP: Q95352; IHK.
DR InterPro: IPRO03006; IG_MHC.
DR InterPro: IPRO03597; IG_C1.
DR InterPro: IPRO01039; MHC_L1.
DR Pfam: PF00128; MHC_1; 1.
DR ProDom: PD000050; MHC_1; 1.
DR SMART: SM00407; IGC1.1; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KM MHC_I; Transmembrane; Glycoprotein; signal.
FT SIGNAL.
FT FT 25 365
FT STAIL.
FT FT 25 365
FT FT 25 365
FT FT DOMAIN 25 114
FT FT DOMAIN 115 206
FT FT DOMAIN 207 298
FT FT DOMAIN 299 308
FT FT TRANSMEM 309 332
FT FT DOMAIN 333 365
FT FT TRANSMEM 366 383
FT FT DISULFID 227 283
FT FT CARBOHYD 110 110
SQ SEQUENCE 365 AA: 40829 MW: D5A0E97410BA1BE CRC64:
Query Match Best Local Similarity 31.9% Score 45; DB 1; Length 365;
Matches 11; Conservative 55; Mismatches 21; Indels 0; Gaps 0.
OY 1 GIGVIXXXXXXWTCGTATGISIKXXXXXXRRQ 37
| | | | | | | | | | | | | | | | | |
310 GIAYLLALCAAVTGAAVGAAYVAAMVRKKSKDKKCSYSQ 346
| | | | | | | | | | | | | | | | | |

```

[illegible]

RESULT 10
 TYPE: LOCAL
 AC MEDLINE-9746615; PubMed-9278503; PRT: 445 AA.
 P45539; P76686; DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-NOV-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transport protein ynfM.
 DE ynfM OR B3370.
 OS Bacillus subtilis.
 OS Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE-9746615; PubMed-9278503;
 RA Riley M., Collado-Vides J., Glaser J.D., Roda C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kikpatrick H.A., Coeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RT Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC -1- SIMILARITY: BELONGS TO THE ABCD/CAB/GRAB/VSI FAMILY OF
 CC PERMEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: U18997; AAC58167.1; ALT_SDB.
 DR EMBL: AE00413; AAC76395.1; ALT_INIT.
 DR ECGene: E312065.ynfM; ynfM.primase1.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00334; aa_permease; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 10 POTENTIAL.
 FT TRANSMEM 36 POTENTIAL.
 FT TRANSMEM 103 POTENTIAL.
 FT TRANSMEM 123 POTENTIAL.
 FT TRANSMEM 135 POTENTIAL.
 FT TRANSMEM 181 POTENTIAL.
 FT TRANSMEM 215 POTENTIAL.
 FT TRANSMEM 236 POTENTIAL.
 FT TRANSMEM 256 POTENTIAL.
 FT TRANSMEM 273 POTENTIAL.
 FT TRANSMEM 334 POTENTIAL.
 FT TRANSMEM 354 POTENTIAL.
 FT TRANSMEM 385 POTENTIAL.
 FT TRANSMEM 411 POTENTIAL.
 FT TRANSMEM 417 POTENTIAL.
 FT TRANSMEM 435 POTENTIAL.
 SQ SEQUENCE 445 AA: 47576 MW: 98805944ASIEPF CRC64:
 Query Match 30.5%; Score 43; DB 1; Length 445;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Caps 0;
 Oy 12 VITGGLPALISMIR 25
 Db 423 VIATGLPAVAFNAK 436
 RESULT 11
 DEGP_RITIME STANDBOARD: PRT: 504 AA.
 ID DEGP_RITIME
 DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE EMBL: U18997; AAC58167.1; ALT_SDB.
 DE P45539; P76686; DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-NOV-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical transport protein ynfM.
 DE ynfM OR B3370.
 OS Bacillus subtilis.
 OS Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RC MEDLINE-9746615; PubMed-9278503;
 RA Riley M., Collado-Vides J., Glaser J.D., Roda C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kikpatrick H.A., Coeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RT Science 277:1453-1474(1997).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC -1- SIMILARITY: BELONGS TO THE ABCD/CAB/GRAB/VSI FAMILY OF
 CC PERMEASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: U18997; AAC58167.1; ALT_SDB.
 DR EMBL: AE00413; AAC76395.1; ALT_INIT.
 DR ECGene: E312065.ynfM; ynfM.primase1.
 DR InterPro: IPR004841; Permease.
 DR Pfam: PF00334; aa_permease; 1.
 KW Hypothetical protein; Transport; Transmembrane; Inner membrane;
 KW Complete proteome.
 FT TRANSMEM 10 POTENTIAL.
 FT TRANSMEM 36 POTENTIAL.
 FT TRANSMEM 103 POTENTIAL.
 FT TRANSMEM 123 POTENTIAL.
 FT TRANSMEM 135 POTENTIAL.
 FT TRANSMEM 181 POTENTIAL.
 FT TRANSMEM 215 POTENTIAL.
 FT TRANSMEM 236 POTENTIAL.
 FT TRANSMEM 256 POTENTIAL.
 FT TRANSMEM 273 POTENTIAL.
 FT TRANSMEM 334 POTENTIAL.
 FT TRANSMEM 354 POTENTIAL.
 FT TRANSMEM 385 POTENTIAL.
 FT TRANSMEM 411 POTENTIAL.
 FT TRANSMEM 417 POTENTIAL.
 FT TRANSMEM 435 POTENTIAL.
 SQ SEQUENCE 445 AA: 47576 MW: 98805944ASIEPF CRC64:
 Query Match 30.5%; Score 43; DB 1; Length 445;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Caps 0;
 Oy 12 VITGGLPALISMIR 25
 Db 423 VIATGLPAVAFNAK 436
 RESULT 11
 DEGP_RITIME STANDBOARD: PRT: 504 AA.
 ID DEGP_RITIME
 DT 30-MAY-2000 (Rel. 39, Created)

SQ SEQUENCE 504 AA: 53055 MM: D7EB2B9981EA23C CRC64:
 Query Match 30.5% Score 43: DB 1: Length 504:
 Percent Similarity 50.0% 12
 Matches 11: Conservative 3: Mismatches 2: Indels 6: Gaps 1:
 Oy 1 GIGAVLKKXXXXVLTGGLPLIS 22
 1: ||| ||||| :
 Db 9 GLAVL-----LTGGLPAEVA 24
 RESULT 12
 CARB_SVTN3
 ID CARB_SVTN3 STANDARD: PRT: 1081 AA.
 AC 055756:
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Carboxymethyl-phosphate synthase large chain (EC 6.3.5.5) (Carboxymethyl-phosphate synthase ammonia chain).
 DE CARB OR SL02370
 OS Synecocystis sp. (Strain PCC 6803).
 OS Bacteria: Cyanobacteria: Chroococcales, Synecocystis.
 NX NCBI_TaxID=1148:
 RA [1]
 RA SEQUENCE FROM N.A.
 RA MEDLINE=97061201; PubMed=9905231;
 RA RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nanno K., Okumura S., Shimo S., Takeuchi C., Wada T., Matsunabe A., Yamada M., Yasuda M., Tabata S.;
 RA "Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp. strain PCC6803. II. Sequence determination of the DNA Res. 3:109-136(1996)."
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP + Phosphate + L-glutamate + carboxymethyl phosphate.
 CC -1- COPROTEIN: Binds three manganese ions (by similarity).
 CC -1- PATHWAY: Arginine biosynthesis: first step.
 CC -1- FUNCTION: Promotes the synthesis of two chains, the small (or glutamine) chain by the large (or ammonia) chain to synthesize carboxymethyl phosphate (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: D64002; BA010403.1; ALT.INT.
 DR HSSP: P00968; 1A98; Cpsae_L.
 DR InterPro: IPR005479; Cpsae_L.D2.
 DR InterPro: IPR005480; Cpsae_L.D3.
 DR InterPro: IPR004362; MGS_Like.
 DR Pfam: PF00289; Cpsae_L.Chain_2.
 DR Pfam: PF02786; Cpsae_L.D2_2.
 DR Pfam: PF02787; Cpsae_L.D3_1.
 DR PRINTS: PR00098; Cpsase.
 DR PROSITE: PS00866; Cpsase_1; 1.
 DR PROSITE: PS00867; Cpsase_2; 2.
 RA Arginine biosynthesis: pyrimidine biosynthesis; Ligase; Repeat:
 RA Arg-binding; Manganese; Complete proteome.
 FT DOMAIN 1 403 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
 FT DOMAIN 404 553 OLIGOMERIZATION DOMAIN.

FT DOMAIN 554 944 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
 FT REPEAT 945 1081 ALLOSTERIC DOMAIN.
 FT REPEAT 554 1081
 FT NP BIND 155 354
 FT NP BIND 155 354
 FT METAL 286 286
 FT METAL 300 300
 FT METAL 302 302
 FT METAL 837 837
 FT METAL 849 849
 SQ SEQUENCE 1081 AA: 119030 MM: OC26053905B40E86 CRC64:
 Query Match 30.1% Score 42.5: DB 1: Length 1081:
 Best Local Similarity 44.0% 2: Mismatches 7: Indels 5: Gaps 1:
 Matches 11: Conservative 4: Mismatches 7: Indels 5: Gaps 1:
 Oy 1 GIGAVLKKXXXXVLTGGLPLIS 25
 1: ||| ||||| :
 Db 996 GVEVWLK-----LHGKPRVLDWIK 1015
 RESULT 13
 NPRT2_HUMAN
 ID NPRT2_HUMAN STANDARD: PRT: 639 AA.
 AC 006495:
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-FEB-2002 (Rel. 41, Last sequence update)
 DT 15-FEB-2002 (Rel. 41, Last annotation update)
 DE Renal sodium-dependent phosphate transporter protein 2 (Sodium/phosphate cotransporter 2) (Na(+)/Pi cotransporter 2) (Renal sodium-phosphate transporter protein 2) (Renal Na(+)-dependent phosphate cotransporter 2).
 DE SL034A1 OR SL017A2 OR NPRT2.
 OS Homo sapiens (Human).
 OS Chordata: Mammalia: Eulacostroni.
 OS Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 NX NCBI_TaxID=9606:
 RA [1]
 RA SEQUENCE FROM N.A.
 RA TISSUE=Kidney.
 RA MEDLINE=93317607; PubMed=8327470;
 RA RA Bilezikian J., Weidner A., Markovitch D., Sorribas V., Stange G., Proc. Natl. Acad. Sci. U.S.A. 90:5979-5983(1993).
 CC *Expression cloning of human and rat renal cortex Na/Pi cotransport. "
 CC -1- FUNCTION: MAY BE INVOLVED IN ACTIVELY TRANSPORTING PHOSPHATE INTO CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: KIDNEY; AND LUNG.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: L13258; AAA63554.1; -
 DR PIR: B48189; B48189.
 DR Gene: HNGC:11019; SL034A1.
 DR MIM: 162309; -
 DR InterPro: IPR003841; Na_Pi_cotrans.
 DR InterPro: IPR003842; Na_Pi_cotrans_1.
 DR Pfam: PF00690; NaPi_cotrans_1.
 DR Transports: Symport; Sodium transport; Transmembrane; Glycoprotein; Phosphorylation.
 FT DOMAIN 1 103 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 104 125 M1 (POTENTIAL).
 FT TRANSMEM 145 145 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 146 163 M2 (POTENTIAL).

```

DR EMBL; M65067; AAA70297.1; -
DR EMBL; M64224; AAA25627.1; -
DR PIR; JH0387; JH0387.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 C:\isb-sib\license\isb-sib.ch).

 DR EMBL: X08036; CAA30831.1; -
 DR PIR: S01697; S01697.
 DR PIR: J00413; J00413.
 DR PIR: S06369; S06369.
 DR HSRP: P07740; 1BRU.
 DR InterPro: IPR002103; Bac_Luciferase.
 DR Pfam: PF00216; Luciferase; 1.
 DR PRINTS: PR00089; BACTERIAL_LUCIFERASE.
 DR PROSITE: PS00434; BACTERIAL_LUCIFERASE.
 KW Photoprotein; Luminescence; Oxidoreductase; Monooxygenase;
 FM Flavoprotein; FMN.
 FT DOMAIN 100 115 REGION OF ACTIVE CENTER-1 (BY
 SIMILARITY)
 FT DOMAIN 279 294 REGION OF ACTIVE CENTER-2 (BY
 SIMILARITY)
 FT SEQUENCE 354 AA: 40370 MW: 569765EB903C7B90 CRC64;
 SQ
 Query Match 29.1%; Score 41; DB 1; Length 354;
 Best Local Similarity 43.8%; Pred. No. 19;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 OY 16 GIPALISWIKXXXXK 31
 |||::|||
 ::
 Db 186 GLPWLWISWITTSKK 201

Search completed: May 9, 2003, 16:16:18
 Job time : 14.2941 secs

THIS PAGE BLANK (OPTO)

[illegible]

```

Query Match      31.2%  Score 44: DB 17: Length 457:
Best Local Similarity 31.8%  Pred. No. 50:
Matches 17: Conservative 5: Mismatches 10: Indels 0: Gaps 0:

OY 3 GAVLKXXXVLTGTGIPALISMI 24
DB 281 GALKKSGSYIMLGLSPFGNMW 302

RESULT 12
OY 018105 PRELIMINARY: PRT: 714 AA.
AC 018105:
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE 723aa, 10 protein.
DE 723aa, 10 protein.
OC Eukaryota; Metazoa; Nematoidea; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OC NCBI_TaxID=6239;
OY 11
RN SEQUENCE FROM N.A.
RA Smye R.:
RA Unpublished (OCT-1996) to the EMBL/GenBank/DBJ databases.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RA *Genome sequence of the nematode C. elegans: A platform for
RT Investigating biology.
RT Science 232:2012-2023(1998).
DE EMBL:AF013271.1
DE Interpro: IPR001003: 77M.chemo2.
DE Pfam: PF01604: 7Tm_L5: 2.
SO SEQUENCE 714 AA: 81848 MW: C3A1E2C1DF4PDBCE CRC64:

Query Match      31.2%  Score 44: DB 5: Length 714:
Best Local Similarity 31.8%  Pred. No. 79:
Matches 9: Conservative 3: Mismatches 10: Indels 0: Gaps 0:

OY 2 IGAVLKXXXVLTGTGIPALISW 23
DB 238 IGAVLQYSPVLVGLTIVASW 259

RESULT 13
OY 09RY43 PRELIMINARY: PRT: 749 AA.
AC 09RY43:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE 714aa, 10 protein.
DE 714aa, 10 protein.
OC Bacteria; Thermus/Delinoococcus group; Delinococci; Delinococcales;
OC Delinococcales; Delinococcus.
OC NCBI_TaxID=1299;
OY 11
RN SEQUENCE FROM N.A.
RA Smith H.O.:
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Ueberlack T., Zaleski C.,
RA Maratheva K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;

```

```

RT *Genome sequence of the radioresistant bacterium Delinococcus
RT radiodurans R1.
RL Science 286:1571-1577(1999).
DR 7883 20048
DR 7883 20048
DR Interpro: IPR000731: HMGCR/patht.57M.
DR Pfam: PF03116: MMLP; 2.
DR PROSITE: PS50156: SSD; 1.
KN Hypothetical protein; Complete proteome.
SO SEQUENCE 749 AA: 79149 MW: 27947ED0E9D144B5 CRC64:

Query Match      31.2%  Score 44: DB 16: Length 749:
Best Local Similarity 30.0%  Pred. No. 83:
Matches 12: Conservative 2: Mismatches 8: Indels 2: Gaps 1:

OY 1 GIGAVLKXXXVLT-TGIPALIS 22
DB 315 GIGAVLVVLTAVLSTPLDALPT 338

RESULT 14
OY 09ATG7 PRELIMINARY: PRT: 762 AA.
AC 09ATG7:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE beta-D-glucosidase.
DE beta-D-glucosidase.
CN CC1756.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OC NCBI_TaxID=155892;
RN SEQUENCE FROM N.A.
RA STRAIN=ATCC 19069 / CR15.
RC MEDLINE=2113698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Fockea J., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Deboy R.T., Nelson W.C., Newton A., Stephens C., Shadock J.B.,
RA Ueberlack T., Tran K., Klotz M., Lamthoran J., Brody A., Berry K.,
RA Ueberlack T., Tran K., Klotz M., Lamthoran J., Brody A., Berry K.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT *Complete genome sequence of Caulobacter crescentus.
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005849; MAK23732.1.
DR TIGR: CC1756.
DR Interpro: IPR001772: GH_3C.
DR Interpro: IPR001772: GH_3C.
DR Interpro: IPR002016: Pectinase.
DR Pfam: PF00933: Glyco_hydro_3.C.1.
DR PRINTS: PR001915: Glyco_hydro_3.C.1.
DR PROSITE: PS00436; PEROXIDASE_2; UNKNOWM_1.
KN Complete proteome.
SO SEQUENCE 762 AA: 81011 MW: C44C6B80D7E02DB CRC64:

Query Match      31.2%  Score 44: DB 16: Length 762:
Best Local Similarity 39.1%  Pred. No. 85:
Matches 9: Conservative 3: Mismatches 11: Indels 0: Gaps 0:

OY 2 IGAVLKXXXVLTGTGIPALISMI 24
DB 516 VSANAKNAVVLGGSPVLPALPL 538

RESULT 15
OY 09RK31 PRELIMINARY: PRT: 2078 AA.
AC 09RK31:
DT 01-OCT-2001 (TREMblrel. 18, Created)

```

DT 01-OCT-2001 (TRENBLURE1.18, Last sequence update)
 DT 01-MAR-2002 (TRENBLURE1.20, Last annotation update)
 CH Mycobacterial protein ml1661.
 CH ML1661
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=470330;
 RX MEDLINE=2108230; PubMed=11214968;
 RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 Takemuchi C., Yamada M., Tabata S.;
 RA Complete genome structure of the nitrogen-fixing symbiotic bacterium
 R. loti.
 RT DNA Res. 7:331-336(2000).
 DR EMBL: AF002998; BAB48983.1; .
 DR Hypothetical protein; Complete proteome.
 SO SEQUENCE 2078 AA: 207618 MW: 235764885.9848 CRC64;
 Query Match 31.2% Score 44; DB 16;
 Identical Similarity 41.8%; Pred. NO. 2,4e+02;
 Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
 QY 1 GIGAVLXXXXXXVLTGTPALI 21
 DB 1627 GAGLGLKGGGTVTTAGTPALV 1647

Search completed: May 9, 2003, 16:17:36
 Job time : 28.0784 secs

AA021737
ID AA021737 standard; Peptide: 26 AA.
AC AA021737;
XX
XX 13-SEP-2002 (first entry)
XX
XX Procytotoxin cytolytic peptide #2.
DE
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
XX
XX ovary; prostate; breast; skin; lung; pancreas; cytolytic.
XX
XX unidentified.
XX
XX Key Location/Qualifiers
XX Modified-site 25 /note "The epsilon amino group of Lys is peptide bonded
XX 26 to the alpha carboxyl group of Phe"
XX Modified-site 26 /note "The alpha carboxyl group of Lys is peptide
XX bonded to the epsilon amino group of Lys"
XX
XX US2002045736-A1.
XX 18-APR-2002.
XX
XX 27-AUG-2001: 2001US-0938623.
XX
XX 09-MAY-2001: 2001US-0851422.
XX
XX (YUXX/) YU X.
XX (MAGN/) WAGNER T E.
XX YU X. Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
XX ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
XX via a peptide bond cleavable by a specific protease
XX
XX Disclosure: Page 7; 21pp: English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
XX bound to an inactivator via a peptide bond, where the peptide bond is
XX susceptible to cleavage by a targeting specific protease. The
XX procytotoxin is used to treat cancer, particularly of the prostate,
XX ovary, breast, skin, lung or pancreas. This sequence represents a
XX procytotoxin cytolytic peptide sequence relating to the invention.
XX
XX Sequence 26 AA:
SO
Query Match 98.4%; Score 126; DB 23; Length 26;
Best Local Similarity 100.0%; Pred. No. 6,4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTVDFGIDKILQLIEDK 25
DB 1 GFATLCTVDFGIDKILQLIEDK 25
RESULT 6
AA021738
ID AA021738 standard; Peptide: 26 AA.
AC AA021738;
XX
XX 13-SEP-2002 (first entry)
XX
XX Procytotoxin cytolytic peptide #3.
DE
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
XX
XX ovary; prostate; breast; skin; lung; pancreas; cytolytic.
XX

XX
XX unidentified.
XX
XX Key Location/Qualifiers
XX Modified-site 17 /note "The epsilon amino group of Lys is peptide bonded
XX 25 to the alpha carboxyl group of Phe"
XX Modified-site 25 /note "The epsilon amino group of Lys is peptide bonded
XX to the alpha carboxyl group of Phe"
XX Modified-site 26 /note "The alpha carboxyl group of Phe is peptide bonded
XX to the epsilon amino group of Lys"
XX
XX US2002045736-A1.
XX 18-APR-2002.
XX
XX 27-AUG-2001: 2001US-0938623.
XX
XX 09-MAY-2001: 2001US-0851422.
XX
XX (YUXX/) YU X.
XX (MAGN/) WAGNER T E.
XX YU X. Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
XX ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
XX via a peptide bond cleavable by a specific protease
XX
XX Disclosure: Page 7; 21pp: English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
XX bound to an inactivator via a peptide bond, where the peptide bond is
XX susceptible to cleavage by a targeting specifically of the prostate,
XX ovary, breast, skin, lung or pancreas. This sequence represents a
XX procytotoxin cytolytic peptide sequence relating to the invention.
XX
XX Sequence 26 AA:
SO
Query Match 98.4%; Score 126; DB 23; Length 26;
Best Local Similarity 100.0%; Pred. No. 6,4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTVDFGIDKILQLIEDK 25
DB 1 GFATLCTVDFGIDKILQLIEDK 25
RESULT 7
AAE18200
ID AAE18200 standard; Peptide: 26 AA.
AC AAE18200;
XX
XX 07-MAY-2002 (first entry)
XX
XX Procytotoxin #1 for treating cancer.
DE
XX Cytotoxic; cytostatic; procytotoxin; epsilon-amino group; cancer; cytostatic;
XX
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX unidentified.
XX
XX Key Location/Qualifiers
XX Modified-site 25 /note "Lys (epsilon-alpha); where (epsilon-alpha)
XX represents a peptide bond between the epsilon amino
XX group of lysine and the alpha carboxyl group of the

FT	XX	adjacent phenylalanine*
XX	PN	
XX	MO200185777-A2.	
XX	15-NOV-2001.	
XX	09-MAY-2001: 2001MO-05A0590.	
XX	09-MAY-2001: 2000US-203063P.	
XX	16-JUN-2000: 2000US-212042P.	
XX		
PA	(GREE-) GREENVILLE HOSPITAL SYSTEM.	
XX		
PT	Yu X. Wagner TE.	
PT	WPI: 2002-154423/20.	
XX		
XX	A procytotoxin comprising a cytotoxic peptide with at least one lysine residue which is converted into a cytotoxin in a target cell is used for treating cancer, particularly of the prostate, skin, ovary or lung	
XX		
PS	Disclousure: Page 12: 33pg: English.	
CC		
CC	The present invention relates to a procytotoxin, comprising a cytotoxic peptide with at least one lysine residue bound by a peptide bond to at least one amino acid via the epsilon-amino group of the lysine residue.	
CC	The procytotoxin is used to treat cancer, particularly of the prostate, ovary, lung, skin, cytolytic peptide also known as channel-forming peptide, apoptosis-inducing peptide, and death inducing peptide and death (apoptosis) upon contact. The present sequence is protoxin.	
CC		
SO	Sequence 26 AA:	
	Query Match 98.4%; Score 126; DB 23; Length 26;	
	Best Local Similarity 100.0%; Pred. No. 6,4e-13;	
	Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GFATLCTKYLDPFQIKLIDLEDK 25	
DB	1 GFATLCTKYLDPFQIKLIDLEDK 25	
	RESULT 8	
	AA021736	
	AA021736 standard; peptide: 27 AA.	
XX		
XX	AA021736:	
XX		
XX	13-SEP-2002 (first entry)	
XX		
XX	Procytotoxin cytolytic peptide #1.	
XX		
XX	Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;	
KW	ovary; prostate; breast; skin; lung; pancreas; cytolytic.	
OS	Unidentified.	
XX		
XX	Key	
FT	Location/Qualifiers	
FT	25 Note: "The epsilon amino group of lys is peptide bonded	
FT	to the gamma carboxyl group of Glu"	
FT	26 Note: "The gamma carboxyl group of this Glu is peptide	
FT	bonded to the epsilon amino group of lys. The alpha amino	
FT	group of this Glu is peptide bonded to the gamma carboxyl	
FT	group of another Glu"	
FT	27 Note: "The gamma carboxyl group of this Glu is peptide	
FT	bonded to the alpha amino group of another Glu"	
XX		
XX	US2002045736-A1.	

PD	18-APR-2002.
PE	27-AUG-2001: 2001US-0938623.
PR	09-MAY-2001: 2001US-0851422.
PS	(YUXX/) YU X
PT	(MAGN/) MAGNER T E.
XX	Yu X, Wagner TE:
PI	WPI: 2002-507251/54.
DR	A new procytotoxin useful in the treatment of cancer of e.g. prostate, ovary, breast, skin, lung or pancreas, this sequence represents a procytotoxin cytotytic peptide sequence relating to the invention.
CC	Disclosure: Page 7; 21pp: English.
XX	The invention relates to a procytotoxin comprising a cytotoxic peptide bound to an inactivator via a peptide bond, where the peptide bond is
CC	procytotoxin is used to treat cancer, particularly of the prostate,
CC	ovary, breast, skin, lung or pancreas. This sequence represents a
CC	procytotoxin cytotytic peptide sequence relating to the invention.
XX	Sequence 27 AA:
QY	Query Match 98.44; Score 126; DB 23; Length 27:
XX	Local similarity: 100.0%; Pctid: NO; 6-7e-13;
XX	Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB	1 GFIATLCYVLPFCPIKGLIOLIEPK 25
DB	1 GFIATLCYVLPFCPIKGLIOLIEPK 25
RESULT 9	
ID	AAE18199 standard; peptide; 27. AA.
XX	XX
AC	AAE18199:
DT	07-MAY-2002 (first entry)
XX	XX
XX	Procytotoxin #3 for treating cancer.
KM	Procytotoxin: cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
OS	Unidentified.
XX	XX
XX	XX
XX	Key
XX	Modified-site
FT	2
FT	/note: "lys-[epsion-gamma]: where [epsion-gamma]
FT	represents a peptide bond between the epsilon amino
FT	group of lysine and the gamma carboxyl group of the
FT	adjacent glutamate"
FT	26
FT	/note: "Glu-[alpha-gamma]: where [alpha-gamma]
FT	represents a peptide bond between the alpha
FT	group of glutamic acid and the gamma carboxyl
FT	group of the second glutamate
XX	XX
XX	WO2001 85777-AZ.
PD	15-NOV-2001.
PD	09-MAY-2001: 2001WO-US040560.
XX	XX
XX	09-MAY-2000: 2000US-203063P.
PR	16-JUN-2000: 2000US-212042P.
XX	XX

PA (GRE-) GREENVILLE HOSPITAL SYSTEM.
 PT Yu X. Wagner TE:
 DR WPI: 2002-154423/20.
 XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
 PT residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 .
 XX Disclosure, Page 11: 33pp: English.
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC (apoptosis) upon contact. The present sequence is procytotoxin.
 XX Sequence 27 AA:
 SO Query Match 98.4%: Score 126; DB 23; Length 27;
 Best Local Similarity 100.0%: Pred. No. 6,7e-13;
 Matches 25: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 GFATLCTKVLDPGIDKLQILIEDK 25
 Db 1 GFATLCTKVLDPGIDKLQILIEDK 25
 RESULT 10
 AMLB201 standard: peptide: 27 AA.
 AC AMLB201:
 XX 07-MAY-2002 (first entry)
 XX Protocolin #2 for treating cancer.
 XX Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
 KM prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
 XX Unidentified.
 OS
 FH Key Location/Qualifiers
 FH Modified-site 11-218
 FT /note: "Lys(epsilon-alpha)-Phe): where [epsilon(alpha)-alpha]
 FT represents a peptide bond between the epsilon amino
 FT group of lysine and the alpha carboxyl group of
 FT phenylalanine, and a standard peptide linkage between
 FT lysine and phenylalanine"
 FT Modified-site 26
 FT /note: "Lys(epsilon-alpha): where [epsilon(alpha)-alpha]
 FT represents a peptide bond between the epsilon amino
 FT group of lysine and the alpha carboxyl group of the
 FT adjacent phenylalanine"
 XX W0200185777-A2.
 PN 15-NOV-2001.
 PD 09-MAY-2001: 2001NO-US40590.
 XX 09-MAY-2000: 2000US-203063P.
 PR 16-JUN-2000: 2000US-212042P.
 XX (GREP-) GREENVILLE HOSPITAL SYSTEM.
 PA Yu X. Wagner TE:
 PT

DR WPI: 2002-154423/20.
 XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
 PT residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 .
 XX Disclosure, Page 12: 33pp: English.
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC (apoptosis) upon contact. The present sequence is procytotoxin.
 XX Sequence 27 AA:
 SO Query Match 90.2%: Score 115.5; DB 23; Length 27;
 Best Local Similarity 96.2%: Pred. No. 3.1e-11;
 Matches 25: Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 Oy 1 GFATLCTKVLDPGIDKLQILIEDK 25
 Db 1 GFATLCTKVLDPGIDKLQILIEDK 26
 RESULT 11
 AMLB201 standard: peptide: 24 AA.
 AC AMLB201:
 XX 13-SEP-2002 (first entry)
 XX Procytotoxin cytolytic peptide #6.
 XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
 KM ovary; prostate; breast; skin; lung; pancreas; cytolytic.
 XX Unidentified.
 OS US2002045736-A1.
 PN 18-APR-2002.
 PD 27-AUG-2001: 2001US-0938623.
 XX 09-MAY-2001: 2001US-0851422.
 PR (YUXX/) YU X.
 PA (WAGN) WAGNER T E.
 XX Yu X. Wagner TE:
 PT WPI: 2002-507251/54.
 XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
 PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator,
 PT via a peptide bond cleavable by a specific protease.
 XX Example 1: Page 10: 21pp: English.
 CC The invention relates to a procytotoxin comprising a cytotoxic peptide
 CC bound to an inactivator via a peptide bond, where the peptide bond is
 CC susceptible to cleavage by a targeting specific protease. The
 CC procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, breast, skin, lung or pancreas. This sequence represents a
 CC procytotoxin cytolytic peptide sequence relating to the invention.
 XX Sequence 24 AA:
 SO

FT the epsilon amino group of lysine and the gamma carboxyl
group of adjacent glutamate, [alpha-gamma] represents a
peptide bond between alpha amino group of glutamate and
gamma carboxyl group of adjacent glutamate
XX MO200185777-AZ.
XX 15-NOV-2001.
XX 09-MAY-2001: 2001MO-US40690.
XX 09-MAY-2000: 2000US-203063P.
XX 16-JUN-2000: 2000US-212042P.
XX (GRIE-) GREENVILLE HOSPITAL SYSTEM.
XX Yu X, Wagner TE;
XX WPI: 2002-154423/20.
XX A procyctoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cycotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX Claim 8; Page 24; 33pp; English.
XX The present invention relates to a procyctoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procyctoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procyctoxin.
XX
XX Sequence 26 AA:
XX
XX Query Match 87.1%; Score 111.5; DB 23; Length 26;
XX Best Local Similarity 96.0%; Pred. No. 1,3e-10;
XX Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX 1 GFIATCTCTVADPGIDMLTQILIDRK 25
XX 1 |||||||||||||||||||
XX DB 1 GFIATCTVADPGIDK-10CILEDK 24
XX
XX RESULT 15
XX AAG06275 standard; Protein: 412 AA.
XX AAG06275:
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 2989.
XX Protein identification: signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
XX EP103405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000: 2000EP-030143P.
XX 25-FEB-1999: 9905-012182Z.
XX 09-MAR-1999: 9905-0123180.
XX 09-MAR-1999: 9905-0122548.
XX 23-MAR-1999: 9905-0122884.
XX 25-MAR-1999: 9905-0120284.
XX

PR 29-MAR-1999: 9905-0126785.
PR 01-APR-1999: 9905-012745Z.
PR 06-APR-1999: 9905-0128214.
PR 15-APR-1999: 9905-0128214.
PR 15-APR-1999: 9905-0129845.
PR 19-APR-1999: 9905-0130077.
PR 21-APR-1999: 9905-0130449.
PR 23-APR-1999: 9905-0130510.
PR 28-APR-1999: 9905-0130891.
PR 30-APR-1999: 9905-0131419.
PR 30-APR-1999: 9905-0132407.
PR 04-MAY-1999: 9905-0132484.
PR 05-MAY-1999: 9905-0132485.
PR 06-MAY-1999: 9905-0132486.
PR 07-MAY-1999: 9905-0132487.
PR 11-MAY-1999: 9905-0132863.
PR 13-MAY-1999: 9905-0134216.
PR 14-MAY-1999: 9905-0134219.
PR 14-MAY-1999: 9905-0134221.
PR 14-MAY-1999: 9905-0134370.
PR 18-MAY-1999: 9905-0134768.
PR 19-MAY-1999: 9905-0134941.
PR 20-MAY-1999: 9905-0135124.
PR 21-MAY-1999: 9905-0135253.
PR 23-MAY-1999: 9905-0135253.
PR 25-MAY-1999: 9905-0136021.
PR 27-MAY-1999: 9905-0136392.
PR 28-MAY-1999: 9905-0136782.
PR 03-JUN-1999: 9905-0137222.
PR 04-JUN-1999: 9905-0137528.
PR 04-JUN-1999: 9905-0137502.
PR 06-JUN-1999: 9905-0138064.
PR 10-JUN-1999: 9905-0138540.
PR 10-JUN-1999: 9905-0138847.
PR 14-JUN-1999: 9905-0139119.
PR 16-JUN-1999: 9905-0139452.
PR 16-JUN-1999: 9905-0139453.
PR 17-JUN-1999: 9905-0139454.
PR 17-JUN-1999: 9905-0139454.
PR 18-JUN-1999: 9905-0139455.
PR 18-JUN-1999: 9905-0139456.
PR 18-JUN-1999: 9905-0139457.
PR 18-JUN-1999: 9905-0139458.
PR 18-JUN-1999: 9905-0139459.
PR 18-JUN-1999: 9905-0139460.
PR 18-JUN-1999: 9905-0139461.
PR 18-JUN-1999: 9905-0139462.
PR 18-JUN-1999: 9905-0139463.
PR 18-JUN-1999: 9905-0139464.
PR 18-JUN-1999: 9905-0139465.
PR 18-JUN-1999: 9905-0139466.
PR 18-JUN-1999: 9905-0139467.
PR 21-JUN-1999: 9905-0139817.
PR 22-JUN-1999: 9905-0139899.
PR 23-JUN-1999: 9905-0140353.
PR 24-JUN-1999: 9905-0140654.
PR 24-JUN-1999: 9905-0140823.
PR 25-JUN-1999: 9905-0140991.
PR 30-JUN-1999: 9905-0141842.
PR 01-JUL-1999: 9905-0142154.
PR 02-JUL-1999: 9905-0142155.
PR 02-JUL-1999: 9905-0142953.
PR 08-JUL-1999: 9905-0142803.
PR 09-JUL-1999: 9905-0142920.
PR 12-JUL-1999: 9905-0142977.
PR 13-JUL-1999: 9905-0143542.
PR 14-JUL-1999: 9905-0143632.
PR 15-JUL-1999: 9905-0144005.
PR 16-JUL-1999: 9905-0144086.
PR 16-JUL-1999: 9905-0144086.

PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144685.
 PR 21-JUL-1999; 99US-0145088.
 PR 21-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145219.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145513.
 PR 27-JUL-1999; 99US-0145519.
 PR 27-JUL-1999; 99US-0145519.
 PR 28-JUL-1999; 99US-0145551.
 PR 28-JUL-1999; 99US-0146386.
 PR 28-JUL-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147202.
 PR 06-AUG-1999; 99US-0147203.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148588.
 PR 13-AUG-1999; 99US-0148588.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 23-AUG-1999; 99US-0149930.
 PR 26-AUG-1999; 99US-0150084.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 07-SEP-1999; 99US-0151439.
 PR 07-SEP-1999; 99US-0151439.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154035.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155143.
 PR 24-SEP-1999; 99US-0155655.
 PR 28-SEP-1999; 99US-0156456.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159253.
 PR 13-OCT-1999; 99US-0159254.
 PR 13-OCT-1999; 99US-0159254.
 PR 13-OCT-1999; 99US-0159255.
 PR 14-OCT-1999; 99US-0159310.
 PR 14-OCT-1999; 99US-0159311.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160762.
 PR 21-OCT-1999; 99US-0160762.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 22-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161922.
 PR 29-OCT-1999; 99US-0162142.

Query Match 42.2% Score 54: DB 21: Length 412:
 Best Local Similarity 58.8%:
 Matches 10: Conservative 3: Mismatches 4: Indels 0: Caps 0:
 QY 9 KVDPGIDKLTQIEDK 25
 DB 249 KADPOVSKLTYAKDK 265

Search completed: May 9, 2003, 16:15:36
 Job time : 39.1176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1995 - 2003 Comugen Ltd.

OK protein - protein search, using sw model

Run on: May 9, 2003, 16:13:10 : Search time 9.79412 Seconds
(without alignments) 81.112 Million cell updates/sec

Title: US-09-851-422a-8

Sequence: 128
1 GFIATLCYVDFGIDKQLQLEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing files 45 summaries

Database :

Issued Patents.AA: *
1: /cgn2-6/pdata1/1/aa/5A.COMB.pep: *
2: /cgn2-6/pdata1/1/aa/5B.COMB.pep: *
3: /cgn2-6/pdata1/1/aa/6A.COMB.pep: *
4: /cgn2-6/pdata1/1/aa/6B.COMB.pep: *
5: /cgn2-6/pdata1/1/aa/6C.COMB.pep: *
6: /cgn2-6/pdata1/1/aa/6D.COMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	39.1	499	4	US-08-426-509A-19
2	50	39.1	499	5	PCT-US95-05008-19
3	49	38.3	272	4	US-08-134-001C-3016
4	49	38.3	272	4	US-08-134-001C-3016
5	47	35.9	432	4	US-08-484-182-7
6	46	35.9	54	2	US-08-237-401A-41
7	46	35.9	279	2	US-08-701-191A-23
8	46	35.9	984	2	US-08-673-789-9
9	46	35.9	984	2	US-08-449-645A-19
10	46	35.9	984	2	US-08-702-367A-19
11	46	35.9	984	5	PCT-US95-04681-19
12	44	34.4	245	4	US-07-834-001C-5640
13	44	34.4	245	4	US-07-834-001C-5640
14	44	34.4	255	2	US-07-857-224B-55
15	44	34.4	271	4	US-09-457-040B-41
16	44	34.4	346	3	US-09-173-581-3
17	44	34.4	346	4	US-09-420-915-3
18	44	34.4	509	3	US-09-039-555B-17
19	44	34.4	509	4	US-08-426-509A-18
20	44	34.4	509	4	US-08-426-509A-18
21	44	34.4	509	4	PCT-US95-05008-19
22	43	33.6	582	4	US-09-134-001C-4730
23	43	33.6	968	4	US-09-228-986-76
24	43	33.6	1064	1	US-08-357-598-2
25	43	33.6	1064	2	US-09-003-289-2
26	43	33.6	1064	5	PCT-US95-16435-2
27	43	33.6	1082	1	US-08-357-598-5

28	43	33.6	1082	1	US-08-357-598-10	Sequence 10, Appl
29	43	33.6	1082	2	US-09-003-289-5	Sequence 5, Appl
30	43	33.6	1082	2	US-09-003-289-10	Sequence 10, Appl
31	43	33.6	1082	3	PCT-US95-16435-10	Sequence 5, Appl
32	43	33.6	1082	3	PCT-US95-16435-10	Sequence 10, Appl
33	43	33.6	1124	4	US-09-191-786-11	Sequence 11, Appl
34	43	33.6	1455	2	US-08-426-509A-2	Sequence 2, Appl
35	42	32.8	675	4	US-08-426-509A-4	Sequence 4, Appl
36	42	32.8	675	5	PCT-US95-05008-4	Sequence 4, Appl
37	42	32.8	766	5	PCT-US94-00198-6	Sequence 6, Appl
38	42	32.8	1098	4	US-08-446-994-17	Sequence 17, Appl
39	42	32.8	1099	4	US-08-665-546-16	Sequence 16, Appl
40	42	32.8	1100	2	US-09-003-289-11	Sequence 11, Appl
41	42	32.8	1100	2	PCT-US95-08354A-2	Sequence 11, Appl
42	42	32.8	1299	5	PCT-US95-16435A-2	Sequence 11, Appl
43	42	32.4	274	2	US-08-701-191A-33	Sequence 2, Appl
44	41.5	32.4	76	2	US-08-469-537A-40	Sequence 30, Appl
45	41	32.0	76	2	US-08-469-537A-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1:
US-08-426-509A-19
Sequence 19, Application US/08426509A
Patent No. 6326469
INVENTOR: IMMAN G.
APPLICANT: GISHIZSKY, Axel
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS: Edmonds
STREET 1155 Avenue of the Americas
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTISO Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-Apr-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: 08/732,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7663-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-869-9741
TELEFAX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
SOURCE: human
MOLBIOL TYPE: No. 6326469e
US-08-426-509A-19

Query Match 39.1%; Score 50; DB 4; Length 499;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 T1CTKVLDFGIDKLI 19
DB 365 T1CTKVLDFGIDKLI 379

RESULT 2

PCT-US95-05008-19
Sequence 19, Application PC/RTUS9505008
GENERAL INFORMATION:
APPLICANT: Bogen, Inc.
APPLICANT: 5150 Madison Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Misesbachten E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80535
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT FILING DATE: 1997-08-14
APPLICATION NUMBER: PCT/US95/05008
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
COUNTRY: US
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEX: 661421863-9741
FAX: 661421863-9741
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULAR WEIGHT: protein
PCT-US95-05008-19

Query Match 39.1% Score 50; DB 5; Length 499;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
OY 5 T1CTKVLDFGIDKLI 19
DB 365 T1CTKVLDFGIDKLI 379

RESULT 3
US-09-134-001C-3016
Sequence 3016, Application US/09134001C
Patent No. 6386370
GENERAL INFORMATION:
APPLICANT: Bryn Doucelte-Stamm et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GPC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1997-08-14
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 3016
TYPE: part
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3016

Query Match 38.3% Score 49; DB 4; Length 273;
Best Local Similarity 43.5%; Pred. No. 5;
Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
OY 1 GFAATCTVTFGASALIPLEK 78
DB 56 GFAATCTVTFGASALIPLEK 78

RESULT 4
US-09-384-162-7
Sequence 7, Application US/09384162
Patent No. 6376747
GENERAL INFORMATION:
APPLICANT: Xing, Ti
APPLICANT: Malik, Kamal
APPLICANT: Malik-Hellier, Teresa
APPLICANT: Malik, Brian
TITLE OF INVENTION: Plant-Derived Map Kinase Kinase
FILE REFERENCE: GB-8428005
CURRENT APPLICATION NUMBER: US/09/384,162
CURRENT FILING DATE: 1999-08-27
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 722
TYPE: part
ORGANISM: Leishmania donovani
US-09-384-162-7

Query Match 36.7% Score 47; DB 4; Length 222;
Best Local Similarity 57.1%; Pred. No. 8;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 9 KVLDFGIDKLI 22
DB 115 KVLDFGIDKLI 128

RESULT 5
US-09-456-647B-41
Sequence 41, Application US/09456647B
Patent No. 5811516
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/08/237/401A
 APPLICATION NUMBER: 02-JUN-1995
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 530
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/237/401
 FILING DATE: 02-MAY-1994
 PRIOR APPLICATION DATA: US 07/884,486
 FILING DATE: 15-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Weherell Ph.D., John R.
 REGISTRATION NUMBER: 31,678
 REFERENCE/DOCKET NUMBER: 07251/007002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 678-5070
 TELEFAX: (619) 678-5099
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 54 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-436-647B-41

Query Match 35.9%; Score 46; DB 2; Length 54;
 Best Local Similarity 57.1%; Pred. No. 2.2;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKYVDFGIDKLI 19
 DB 8 LCKKVSDFGLRL 21

RESULT 6
 US-08-237-401A-41
 Sequence 41, Application US/08237401A
 Patent No. 5637448
 GENERAL INFORMATION:
 INVENTOR: Ph.D. et al., Greg E.
 TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
 NUMBER OF SEQUENCES: 54
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: US
 ZIP: 92037

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/237/401A
 FILING DATE: 02-MAY-1994
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/884,486
 FILING DATE: 15-MAY-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Weherell Ph.D., John R.
 REGISTRATION NUMBER: 31,678
 REFERENCE/DOCKET NUMBER: 07251/007001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 678-5070
 TELEFAX: (619) 678-5099
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:

LENGTH: 54 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-237-401A-41

Query Match 35.9%; Score 46; DB 2; Length 54;
 Best Local Similarity 57.1%; Pred. No. 2.2;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKYVDFGIDKLI 19
 DB 8 LCKKVSDFGLRL 21

RESULT 7
 US-08-701-191A-23
 Sequence 23, Application US/08701191A
 Patent No. 5637448
 GENERAL INFORMATION:
 INVENTOR: Moosa Mohammadi, Joseph Schlössinger,
 Applicant: and Steven R. Hubbard
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
 OF NON-INSULIN RECEPTOR TYROSINE KINASE
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: 613 Avenue of the
 CITY: Los Angeles
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 90071-2066

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 Diskette, 1.44 MB
 COMPUTER: 386
 OPERATING SYSTEM: IBM compatible
 SOFTWARE: FastISO for Windows 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/701,191A
 FILING DATE: August 21, 1996
 CLASSIFICATION: 350
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Warburg, Richard J.
 REGISTRATION NUMBER: 32,327
 REFERENCE/DOCKET NUMBER: 827/088
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 955-0400
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 279 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-701-191A-23

Query Match 35.9%; Score 46; DB 2; Length 279;
 Best Local Similarity 57.1%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKYVDFGIDKLI 19
 DB 151 LCKKVSDFGLRL 164

RESULT 8
US-08-673-789-9
Sequence 9: Application US/08673789
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: EHOW, RENNING, SCHULZ, NICHOLAS,
APPLICANT: T.: KROMER, LAMERCE, F.: VANDE NOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: KINASES
TITLE OF INVENTION: METHOD IN DIAGNOSTIC AND THERAPEUTIC
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
FILING DATE: 09/08/673,789
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. CRUPP
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800 ..
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 984
TYPE: AMINO ACID
STANDARDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-9
Query Match 35.9%, Score 46; DB 2; Length 984;
Best Local Similarity 57.1%, Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Caps 0:
DB 769 LCCVNSDFLTRLL 782
OY 6 LCTKYVLDPCIDKLI 19
II III III: II:
RESULT 9
US-08-449-645A-19
Sequence 19: Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-like Receptor Protein Tyrosine
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Patent Operations/RBW
STREET: 1840 Delavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. CRUPP
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 amino acids
TYPE: amino acid
STANDARDNESS: single
TOPOLOGY: linear
MOLBULE TYPE: protein
US-08-449-645A-19
Query Match 35.9%, Score 46; DB 2; Length 984;
Best Local Similarity 57.1%, Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Caps 0:
OY 6 LCTKYVLDPCIDKLI 19
II III III: II:
DB 769 LCCVNSDFLTRLL 782
RESULT 10
US-08-702-367A-19
Sequence 19: Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: Eph-like Receptor Protein Tyrosine
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Angen Patent Operations/RBW
STREET: 1840 Delavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. CRUPP
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 amino acids
TYPE: amino acid
STANDARDNESS: single
TOPOLOGY: linear
MOLBULE TYPE: protein
US-08-702-367A-19
Query Match 35.9%, Score 46; DB 2; Length 984;
Best Local Similarity 57.1%, Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Caps 0:
OY 6 LCTKYVLDPCIDKLI 19
II III III: II:

Db 769 LCCXVSDPGLRLL 782

```

RESULT 11
PCT-US95-04681-19
: Sequence 19, Application PC/TUS9504681
: GENERAL INFORMATION:
: APPLICANT: Fox, Gary M.
: TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
: NUMBER OF INVENTIONS: 28
: NUMBER OF SEQUENCES: 28
: ADDRESS: 10000 Wilshire Blvd, Suite 1000, Beverly Hills, CA 90210
: ADDRESSER: Amgen Patient Operations/BNM
: STREET: 1840 Dehaveland Drive
: CITY: Thousand Oaks
: STATE: California
: COUNTRY: USA
: ZIP: 91320
: COMPUTER READABLE FORM:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04681
: FILING DATE:
: PRIORITY DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Milner, Robert B.
: REFERENCE/DOCKET NUMBER: A-287
: INFORMATION FOR SEQ ID NO: 19:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 984 amino acids
: TYPE: amino acid
: STRATEGY: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
PCT-US95-04681-19

```

Query Match 35.9%; Score 46; DB 5; Length 984;
 Best Local Similarity 57.1%; Pred. No. 64;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTVLDPRDKLI 19
 Db 769 LCCXVSDPGLRLL 782

```

RESULT 12
US-09-134-001C-3640
: Sequence 3640, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: CURRENT INVENTION NUMBER: US/09/134, 001C
: CURRENT FILING DATE: 1998-06-13
: PRIOR APPLICATION NUMBER: US 60/054,564
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO: 3640
: LENGTH: 148
: TYPE: prt
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3640

```

Query Match 34.4%; Score 44; DB 4; Length 148;
 Best Local Similarity 53.8%; Pred. No. 14;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

OY 12 DPGIDKLIQIED 24
 Db 48 DPGIDVAVAIKID 60

```

RESULT 13
US-07-857-224B-54
: Sequence 54, Application US/07857224B
: Patent No. 5958784
: GENERAL INFORMATION:
: APPLICANT: Banner, Steven A.
: TITLE OF INVENTION: Predicting Folded Structures of Proteins
: NUMBER OF INVENTIONS: 114
: CORRESPONDENCE ADDRESS:
: ADDRESSER: Steven A. Banner
: STREET: Hadlaubstrasse 151
: CITY: Zurich
: STATE: none
: COUNTRY: Switzerland
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
: COMPUTER: Apple Macintosh
: OPERATING SYSTEM: Macintosh 7.0
: SOFTWARE: Microsoft Word
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/857,224B
: FILING DATE: 03/25/92
: CLASSIFICATION: 436
: PRIOR APPLICATION DATA:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (international) 41 1 632 2830
: TELEFAX: (international) 41 1 262 2437
: TEXT: none
: INFORMATION FOR SEQ ID NO: 54:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 259
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGIN: human
: ORGANISM: human
: FEATURE: Protein kinase: Table 8 column 61
: PUBLICATION INFORMATION:
: AUTHORS:
: AUTHORS: Hanks, S. K.
: AUTHORS: Hunter, T.
: AUTHORS: Quinn, A. M.
: JOURNAL: The protein kinase family
: VOLUME: 241
: PAGES: 42-52
: DATE: 1998
US-07-857-224B-54

```

Query Match 34.4%; Score 44; DB 2; Length 259;
 Best Local Similarity 50.6%; Pred. No. 29;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 TCTVLDPRDKLI 20
 Db 133 TUSCKIADPGLRLL 148

Query Match 34.4%; Score 44; DB 2; Length 259;
 Best Local Similarity 50.6%; Pred. No. 29;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

US-09-457-040B-41
Query Match      34.48: Score 44: DB 4: Length 271:
Best Local Similarity 50.0%: Pred. No. 29:
Matches      8: Conservative      4: Mismatches      4: Indels      0: Caps      0:
OY      5 TCTCXYADPDKIMK 20
      || | | | | | | | |
Db      145 TLTCKINDPOLARITE 160

Search completed: May 9, 2003, 16:19:29
Job time : 11.7941 secs

```

GenCore version 5.1.4-p5-4578
Copyright (c) 1995 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:15:46 : Search time 11.3824 Seconds

(without alignments)
218.294 Million cell updates/sec

Title: US-09-851-422A-8

Sequence: 128
1 GFIALCTGVDFGIDKLIQLERXXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Published Applications, AA:*

1: /cgn2_6/p/ctodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/p/ctodata/2/pubpaa/PC1_NEW_PUB.pep:*
3: /cgn2_6/p/ctodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/p/ctodata/2/pubpaa/US05_PUBCOMB.pep:*
5: /cgn2_6/p/ctodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/p/ctodata/2/pubpaa/PC1US_PUBCOMB.pep:*
7: /cgn2_6/p/ctodata/2/pubpaa/PC1US_PUBCOMB.pep:*
8: /cgn2_6/p/ctodata/2/pubpaa/US08_NEW_PUB.pep:*
9: /cgn2_6/p/ctodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/p/ctodata/2/pubpaa/US10_NEW_PUB.pep:*
11: /cgn2_6/p/ctodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/p/ctodata/2/pubpaa/US10_NEW_PUB.pep:*
13: /cgn2_6/p/ctodata/2/pubpaa/US10_NEW_PUB.pep:*
14: /cgn2_6/p/ctodata/2/pubpaa/US06_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	50	38.1	499	9	US-09-977-260-19
2	50	38.1	499	9	US-09-977-260-19
3	50	38.1	499	9	US-09-977-260-19
4	50	38.1	499	9	US-09-977-260-19
5	46	35.9	531	10	US-09-158-722-41
6	46	35.9	531	10	US-09-158-722-41
7	46	35.9	531	10	US-09-158-722-41
8	46	35.9	531	10	US-09-158-722-41
9	46	35.9	531	10	US-09-158-722-41
10	46	35.9	531	10	US-09-158-722-41
11	46	35.9	531	10	US-09-158-722-41
12	46	35.9	531	10	US-09-158-722-41
13	46	35.9	531	10	US-09-158-722-41
14	46	35.9	531	10	US-09-158-722-41
15	46	35.9	531	10	US-09-158-722-41
16	46	35.9	531	10	US-09-158-722-41
17	46	35.9	531	10	US-09-158-722-41
18	46	35.9	531	10	US-09-158-722-41
19	46	35.9	531	10	US-09-158-722-41
20	46	35.9	531	10	US-09-158-722-41

20	43	33.6	169	9	US-10-101-464A-562	Sequence 562, App
21	43	33.6	309	9	US-10-101-464A-602	Sequence 602, App
22	43	33.6	347	10	US-09-815-242-10754	Sequence 10754, A
23	43	33.6	674	9	US-10-086-164-115	Sequence 115, App
24	43	33.6	674	9	US-10-086-164-115	Sequence 115, App
25	43	33.6	674	9	US-10-086-164-115	Sequence 115, App
26	43	33.6	674	9	US-10-086-164-115	Sequence 115, App
27	42.5	33.2	423	9	US-09-738-526-6367	Sequence 6367, App
28	42	32.8	172	10	US-09-810-264-26	Sequence 26, App
29	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
30	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
31	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
32	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
33	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
34	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
35	42	32.8	626	9	US-09-712-563-191	Sequence 191, App
36	41	32.0	188	9	US-09-738-526-5200	Sequence 5200, App
37	41	32.0	505	9	US-09-977-260-17	Sequence 17, App
38	41	32.0	505	9	US-09-977-260-17	Sequence 17, App
39	41	32.0	505	9	US-09-977-260-17	Sequence 17, App
40	41	32.0	505	9	US-09-977-260-17	Sequence 17, App
41	41	32.0	3672	10	US-09-927-666-10	Sequence 10, App
42	41	32.0	3672	10	US-09-927-666-10	Sequence 10, App
43	40.5	31.6	192	9	US-10-101-482-13	Sequence 13, App
44	40.5	31.6	192	9	US-10-101-482-13	Sequence 13, App
45	40	31.2	119	10	US-09-815-242-11367	Sequence 11367, A

ALIGNMENTS

RESULT 1
US-09-977-260-19
Sequence 19, Appl
Publication No. US002019279041
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIKY, MIKAHI
TITLE OF INVENTION: NOVEL MEANAROCYTIC PROTEIN TYROSINE KINASES
CURRENT FILING DATE: 1994-04-22
PRIOR APPLICATION NUMBER: 08/232,545
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentia Ver. 2.1
SEQ ID NO 19
LENGTH: 499
ORIGIN: Mus sp.
US-09-977-260-19

Query Match 39.1%: Score 50; DB 9; Length 499;
Best Local Similarity 53.3%: Pred. No. 8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

5 TCTCTVDFGIDKLI 19
DB 365 TCTCTVDFGIDKLI 379

RESULT 2
US-09-977-261-19
Sequence 19, Appl
Publication No. US002019279041
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIKY, MIKAHI
TITLE OF INVENTION: NOVEL MEANAROCYTIC PROTEIN TYROSINE KINASES
CURRENT FILING DATE: 1994-04-22
PRIOR APPLICATION NUMBER: 08/232,545
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentia Ver. 2.1
SEQ ID NO 19
LENGTH: 499
ORIGIN: Mus sp.
US-09-977-261-19

```

? CURRENT FILING DATE: 2001-10-16
? PRIOR APPLICATION NUMBER: 08/232,545
? PRIOR FILING DATE: 1994-04-22
? NUMBER OF SEQ ID NOS: 2
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 19
? LENGTH: 499
? TYPE: PRT
? ORGANISM: Mus sp.
US-09-977-261-19
Query Match          39.1% Score 50; DB 9; Length 499;
Best Local Similarity 53.3% Pred. No. 8.8;
Matches      8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY      5 TICTKVLDFGIDKLI 19
DB      365 TICTKVLDFGIDKLI 379

RESULT 3
US-09-977-269-19
? Sequence 19, Application US/09977269
? Patent No. US20020082037A1
? GENERAL INFORMATION:
? ORGANISM: AXEL
? APPLICANT: GSHUZYX/ MIKHAIL
? APPLICANT: SURES, IRMINGARD
? TITLE OF INVENTION: NOVEL MEKANAROCYTIC PROTEIN TYROSINE KINASES
? FILE REFERENCE: 038602/1260
? CURRENT APPLICATION NUMBER: US/09/977 269
? CURRENT FILING DATE: 2001-10-16
? PRIOR FILING DATE: 1994-04-22
? NUMBER OF SEQ ID NOS: 24
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 19
? LENGTH: 499
? TYPE: PRT
? ORGANISM: Mus sp.
US-09-977-269-19
Query Match          39.1% Score 50; DB 10; Length 499;
Best Local Similarity 53.3% Pred. No. 8.8;
Matches      4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY      5 TICTKVLDFGIDKLI 19
DB      365 TICTKVLDFGIDKLI 379

RESULT 4
US-10-101-464A-492
? Sequence 492, Application US/10101464A
? Publication No. US20030046728A1
? ORGANISM: Neuwahlzahn, Timothy
? APPLICANT: Stachelin, Nicholas
? APPLICANT: Hinglins, Colleen M.
? TITLE OF INVENTION: Compositions Isolated from Plant Cells
? TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
? FILE REFERENCE: 11000.1020c2
? CURRENT APPLICATION NUMBER: US/10/101,464A
? CURRENT FILING DATE: 2000-11-01
? PRIOR FILING DATE: 1999-01-12
? PRIOR FILING DATE: 1999-01-12
? PRIOR FILING DATE: 1999-01-12
? PRIOR FILING DATE: 1999-11-01
? PRIOR APPLICATION NUMBER: PC9/US00/00724
? PRIOR FILING DATE: 2000-11-11
? NUMBER OF SEQ ID NOS: 589

```

```

? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 492
? LENGTH: 499
? TYPE: PRT
? ORGANISM: Encalyptus grandis
US-10-101-464A-492
Query Match          38.3% Score 49; DB 9; Length 205;
Best Local Similarity 56.2% Pred. No. 4.6;
Matches      5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY      4 ATLCTVLDGIDKLI 19
DB      86 ABECPVADGIAKIV 101

RESULT 5
US-09-158-722-41
? Sequence 41, Application US/09158722
? Publication No. US20030013848A1
? GENERAL INFORMATION:
? APPLICANT: Leake Ph.D. et al., Greg E.
? TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
? NUMBER OF SEQUENCES: 54
? ORGANISM: FISH 6
? ADDRESS: Fish 6 Richardson P.C.
? STREET: 4225 Executive Square, Suite 1400
? CITY: La Jolla
? STATE: CA
? COUNTRY: US
? ZIP: 92037
? COMPUTER TYPE: FLOPPY disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/158,722
? FILING DATE:
? PRIORITY INFORMATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/456,647
? FILING DATE: 02-JUN-1995
? APPLICATION NUMBER: US 08/237,401
? FILING DATE: 02-MAY-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/884,486
? FILING DATE: 02-MAY-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: McHargill Ph.D., John R.
? REGISTRATION NUMBER: 31,678
? TELEPHONE: (619) 678-5070
? TELEFAX: (619) 678-5070
? INFORMATION FOR SEQ ID NO: 41:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 54 amino acids
? TYPE: amino acid
? STRANDEDNESS: not relevant
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-09-158-722-41
Query Match          35.9% Score 46; DB 9; Length 54;
Best Local Similarity 57.1% Pred. No. 2.9;
Matches      8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY      6 LCTVLDGIDKLI 19
DB      8 LCKVSDGIDKLI 21

```

```

RESULT 6
US-09-881-752a-326
? Sequence 326, Application US/09801752a
? Patent No. US20010130784a1
? GENERAL INFORMATION:
? APPLICANT: Kleenhouss, Harold
? APPLICANT: Al-Gareyl, Amal
? APPLICANT: Miller, Charles
? APPLICANT: Tomb, Jean-Francois
? APPLICANT: Oomen, Raymond P.
? TITLE OF INVENTION: Identification of Polynucleotides
? TITLE OF INVENTION: Genocoding No. US20010130784a1 Helicobacter Polypeptides in the
? FILE REFERENCE: 06132/041002
? CURRENT APPLICATION NUMBER: US/09/881,752A
? PRIOR FILING DATE: 2001-06-15
? PRIOR APPLICATION NUMBER: US 08/833,457
? PRIOR FILING DATE: 1997-04-01
? NUMBER OF SEQ ID NOS: 370
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 326
? LENGTH: 131
? TYPE: PRT
? ORGANISM: Helicobacter pylori
US-09-881-752a-326

Query Match          35.9%  Score 46:  DB 10:  Length 131:
Best Local Similarity 53.34%  Pred. No. 8:
Matches 8:  Conservative 5;  Mismatches 2;  Indels 0;  Gaps 0;

Db 20  LQYVDFGDKLJOLIEDK 25
||||| : : : : :

RESULT 7
US-10-101-464a-729
? Sequence 729, Application US/10101464a
? Publication No. US20030046728a1
? GENERAL INFORMATION:
? APPLICANT: Strabala, Timothy
? APPLICANT: Muehlenetzen, Nicholas
? APPLICANT: Hirsch, Catherine
? TITLE OF INVENTION: Compositions Isolated from Plant Cells
? FILE REFERENCE: 11000.1020c2
? CURRENT APPLICATION NUMBER: US/10/101,464A
? CURRENT FILING DATE: 2002-03-18
? PRIOR APPLICATION NUMBER: 09/704,302
? PRIOR FILING DATE: 2000-11-01
? PRIOR FILING DATE: 1998-01-12
? PRIOR FILING DATE: 1998-01-12
? PRIOR APPLICATION NUMBER: 60/162,866
? PRIOR FILING DATE: 1999-11-01
? PRIOR APPLICATION NUMBER: PCT/US00/00724
? PRIOR FILING DATE: 2000-01-11
? NUMBER OF SEQ ID NOS: 989
? SOFTWARE: FASTSEQ for Windows Version 4.0
? SEQ ID NO 729
? LENGTH: 161
? TYPE: PRT
? ORGANISM: Pinus radiata
US-10-101-464a-729

Query Match          35.9%  Score 46:  DB 9:  Length 161:
Best Local Similarity 53.0%  Pred. No. 10:
Matches 8:  Conservative 3;  Mismatches 5;  Indels 0;  Gaps 0;

Db 85  SNRCPKVMDFGLAKLL 100
| | | | | : : : : :

```

```

RESULT 8
US-09-801-368-376
? Sequence 376, Application US/09801368
? Patent No. US20010128250a1
? GENERAL INFORMATION:
? APPLICANT: Bushy, Robert
? APPLICANT: Call, Brian
? APPLICANT: Hecht, Peter
? APPLICANT: Holtzman, Doug
? APPLICANT: Madden, Kevin
? APPLICANT: Naxon, Mary
? APPLICANT: Nine, Todd
? APPLICANT: Rober, John
? APPLICANT: Salama, Sofie
? APPLICANT: Sherman, Amir
? APPLICANT: Silva, Jeff
? APPLICANT: Summers, Eric
? TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
? FILE REFERENCE: 10972.148
? CURRENT APPLICATION NUMBER: US/09/801,368
? CURRENT FILING DATE: 2001-03-07
? PRIOR APPLICATION NUMBER: US 09/487,558
? PRIOR FILING DATE: 2000-01-19
? PRIOR APPLICATION NUMBER: US 60/160,587
? PRIOR FILING DATE: 1999-10-20
? NUMBER OF SEQ ID NOS: 440
? SOFTWARE: PatentIn Version 3.0
? SEQ ID NO 376
? LENGTH: 717
? TYPE: PRT
? ORGANISM: Saccharomyces cerevisiae
US-09-801-368-376

Query Match          35.9%  Score 46:  DB 10:  Length 717:
Best Local Similarity 42.1%  Pred. No. 58:
Matches 8:  Conservative 3;  Mismatches 8;  Indels 0;  Gaps 0;

Db 592  CVKTIIDGSKSLPLNKK 610
| | | | | : : : : :

RESULT 9
US-05-771-161a-95
? Sequence 95, Application US/09771161a
? Patent No. US2002010811a1
? GENERAL INFORMATION:
? APPLICANT: Levine, et al.
? TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
? FILE REFERENCE: 802920-2005.1
? CURRENT APPLICATION NUMBER: US/05/771,161A
? CURRENT FILING DATE: 2001-01-26
? PRIOR APPLICATION NUMBER: 09/774,676
? PRIOR FILING DATE: 2000-11-28
? PRIOR APPLICATION NUMBER: 136776
? PRIOR FILING DATE: 2000-06-15
? PRIOR APPLICATION NUMBER: 135619
? PRIOR FILING DATE: 2000-04-12
? PRIOR FILING DATE: 2000-07-29
? SOFTWARE: PatentIn Version 3.0
? SEQ ID NO 95
? LENGTH: 454
? TYPE: PRT
? ORGANISM: Homo sapiens
US-05-771-161a-95

Query Match          35.2%  Score 45:  DB 10:  Length 454:
Best Local Similarity 50.0%  Pred. No. 48:
Matches 7:  Conservative 4;  Mismatches 3;  Indels 0;  Gaps 0;

Db 6  LCTKYVDFGDKLKI 19
| | | | | : : : : :

```

Db 321 LCKRIADPGIARI 334

RESULT 10

US-05-741-669-299

Sequence 299, Application US/09741669

Patent No. US20020022718A1

GENERAL INFORMATION:

APPLICANT: Forsyth, R. Allyn

APPLICANT: Ohlson, Karl L.

APPLICANT: Zyskind, Judith W.

TITLE OF INVENTION: Genes identified as required for protein synthesis in the regulation of E. coli

CURRENT FILING DATE: 2000-12-19

PRIOR FILING DATE: 1999-12-23

NUMBER OF SEQ ID NOS: 481

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 299

LENGTH: 485

TYPE: PRT

ORGANISM: Escherichia coli

US-09-741-669-299

Query Match 35.2% Score 45, DB 10, Length 485.

Best Local Similarity 40.9% Pred. No. 51

Matches 3: Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 1 GATATCTCTVLDPGIDKILQIL 22

Db 240 GYIASLCLCWDFFSMRLPLI 261

RESULT 11

US-09-771-161A-186

Sequence 186, Application US/09771161A

Patent No. US20020110811A1

GENERAL INFORMATION:

APPLICANT: LEVINE, et al.

TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES

CITE REFERENCE: 802620-2005.1

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: 09/774,676

PRIOR FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 136776

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: 135619

PRIOR FILING DATE: 2000-04-12

SOFTWARE: Patentin Version 3.0

SEQ ID NO 186

LENGTH: 505

TYPE: PRT

ORGANISM: Homo sapiens

US-09-771-161A-186

Query Match 35.2% Score 45, DB 10, Length 505.

Best Local Similarity 50.0% Pred. No. 54

Matches 7: Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 6 LCKRYLDFGIDKIL 19

Db 372 LCKRIADPGIARI 385

RESULT 12

US-10-101-464A-915

Sequence 915, Application US/10101464A

Patent No. US20030046728A1

GENERAL INFORMATION:

APPLICANT: Strubala, Timothy

APPLICANT: Nieuwenhuizen, Nicolaas

TITLE OF INVENTION: Compositions isolated from Plant Cells

CITE REFERENCE: 101000-102000

CURRENT APPLICATION NUMBER: US/10/101,464A

CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/704,302

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 09/228,986

PRIOR FILING DATE: 1999-01-12

PRIOR FILING DATE: 1998-11-01/62,866

PRIOR APPLICATION NUMBER: PCT/US00/00724

PRIOR FILING DATE: 2000-01-11

NUMBER OF SEQ ID NOS: 989

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 915

LENGTH: 827

TYPE: PRT

ORGANISM: Eucalyptus grandis

US-10-101-464A-915

Query Match 35.2% Score 45; DB 9; Length 827.

Best Local Similarity 50.0% Pred. No. 94

Matches 8: Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATCTKTVLDPGIDKIL 19

Db 629 AETCEPADGICATLY 644

RESULT 13

US-10-101-464A-712

Sequence 712, Application US/10101464A

Patent No. US20030046728A1

GENERAL INFORMATION:

APPLICANT: Strubala, Timothy

APPLICANT: Nieuwenhuizen, Nicolaas

TITLE OF INVENTION: Compositions isolated from Plant Cells

CITE REFERENCE: 101000-102000

CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/704,302

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 09/228,986

PRIOR FILING DATE: 1999-01-12

PRIOR APPLICATION NUMBER: 60/162,866

PRIOR FILING DATE: 1998-11-01/62,866

PRIOR APPLICATION NUMBER: PCT/US00/00724

PRIOR FILING DATE: 2000-01-11

NUMBER OF SEQ ID NOS: 989

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 712

LENGTH: 122

TYPE: PRT

ORGANISM: Pinus radiata

US-10-101-464A-712

Query Match 34.4% Score 44; DB 9; Length 122.

Best Local Similarity 50.0% Pred. No. 15

Matches 10: Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 ATCTKTVLDPGIDKIL 23

Db 90 ADEKATVADVAVALLQTRDE 109

RESULT 14

US-05-870-962-3

Sequence 962, Application US/0505870962A1

Patent No. US20020022718A1

GENERAL INFORMATION:

APPLICANT: Strubala, Timothy

APPLICANT: Nieuwenhuizen, Nicolaas

TITLE OF INVENTION: Compositions isolated from Plant Cells

CITE REFERENCE: 101000-102000

CURRENT APPLICATION NUMBER: US/05/058,962A1

CURRENT FILING DATE: 2002-03-18

PRIOR APPLICATION NUMBER: 09/704,302

PRIOR FILING DATE: 2000-11-01

PRIOR APPLICATION NUMBER: 09/228,986

PRIOR FILING DATE: 1999-01-12

PRIOR FILING DATE: 1998-11-01/62,866

PRIOR APPLICATION NUMBER: PCT/US00/00724

PRIOR FILING DATE: 2000-01-11

NUMBER OF SEQ ID NOS: 989

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 962

LENGTH: 827

TYPE: PRT

ORGANISM: Eucalyptus grandis

US-05-870-962-3

Query Match 35.2% Score 45; DB 9; Length 827.

Best Local Similarity 50.0% Pred. No. 94

Matches 8: Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 4 ATCTKTVLDPGIDKIL 19

Db 629 AETCEPADGICATLY 644

Sequence 3, Application US/09870962
 Patent No. US20020081290A1
 GENERAL INFORMATION: OJBA
 APPLICANT: Bessman, Y. Tosa
 APPLICANT: Rong, Y. Tosa
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Cortley, Neil C.
 APPLICANT: Goergene, Gina
 APPLICANT: Azimzali, Yalda
 APPLICANT: Kozlov, Alexander
 TITLE OF INVENTION: Protein Kinase Homologs
 FILE REFERENCE: PP-0614 US
 CURRENT APPLICATION NUMBER: US/09/870,962
 CURRENT FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: 09/420,915
 PRIOR FILING DATE: 1999-10-20/9/173,561
 PRIOR FILING DATE: 1999-09-15
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PERL Program
 SEQ ID NO 3
 LENGTH: 346
 TYPE: PRT
 ORGANISM: Homo sapiens
 OTHER INFORMATION: 507669
 US-09-870-962-3

Query Match 34.4%, Score 44; DB 10; Length 346;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 5 TICTKYIAPGIDKLTQ 20
 |||: |||: |||:
 Db 212 TISCKIADPGIARLIE 227

RESULT 15
 US-09-977-260-18
 Sequence 18, Application US/09977260
 Patent No. US20020132790A1
 GENERAL INFORMATION:
 APPLICANT: ULLRICH, AXEL
 APPLICANT: GISHIZKY, MIKHAIL
 APPLICANT: SURES, IRMINGARD
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 FILE REFERENCE: 038605/1260 US/09/977,260
 CURRENT APPLICATION NUMBER: US/09/977,260
 CURRENT FILING DATE: 2001-10-14
 PRIOR APPLICATION NUMBER: 08/232,545
 PRIOR FILING DATE: 1994-04-22
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 18
 LENGTH: 509
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-977-260-18

Query Match 34.4%, Score 44; DB 9; Length 509;
 Best Local Similarity 50.0%; Pred. No. 77;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 5 TICTKYIAPGIDKLTQ 20
 |||: |||: |||:
 Db 375 TISCKIADPGIARLIE 390

Search completed: May 9, 2003, 16:20:18
 Job time : 12.3824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: May 9, 2003, 16:08:20 : Search time 8.73529 Seconds

(without alignments) 128.200 Million cell updates/sec

Title: US-09-851-422a-8

Sequence: 1 GFIALCTKVDGIDKIDLEKXX 27

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : swissprot_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	98.4	98	1 PFPA_ENTMH	P34095 entamoeba h
2	125	97.9	97	1 PFPA_ENTMH	O07831 entamoeba h
3	125	97.9	97	1 PFPA_ENTMH	O24824 entamoeba h
4	76	59.4	101	1 PFPA_ENTMH	O97843 syncytiotrophoblast
5	54	42.2	521	1 SPKA_SYN3	P03299 varicella-z
6	53	41.4	835	1 OBP_YZVD	P16277 mus musculus
7	50	39.1	498	1 BLK_MOUSE	P42683 gallus gall
8	49	38.3	507	1 LCK_CHICK	O92949 hellicobacte
9	48.5	37.9	559	1 PRIM_HELLO	P56004 icellulid h
10	47.5	37.1	559	1 PRIM_HELLO	O58130 icellulid h
11	46	35.9	535	1 GGP_METHA	P23561 saccharomy
12	46	35.9	535	1 GGP_METHA	P23561 saccharomy
13	46	35.9	535	1 GGP_METHA	P23561 saccharomy
14	46	35.9	535	1 GGP_METHA	P23561 saccharomy
15	45.5	35.5	660	1 EPAL_HUMAN	P73207 mycoplasma
16	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
17	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
18	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
19	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
20	44.5	34.8	441	1 Y390_MYCN	P54956 bacillus su
21	44.5	34.8	441	1 Y390_MYCN	P54956 bacillus su
22	44	34.4	261	1 YGSI_YEAST	O87411 bruceella me
23	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
24	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
25	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
26	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
27	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
28	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
29	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
30	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
31	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
32	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
33	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me

Result No.	Score	Query Match	Length	ID	Description
34	42	32.8	163	1 ATRP_OCHNE	O40508 ochrosphaer
35	42	32.8	395	1 MENC_ECOLI	P06721 escherichia
36	42	32.8	503	1 HCK_BAT	P05045 rattus norv
37	42	32.8	504	1 HCK_MOUSE	O95m30 mecaea fasc
38	42	32.8	504	1 HCK_MOUSE	O95m30 mecaea fasc
39	42	32.8	504	1 HCK_MOUSE	O95m30 mecaea fasc
40	42	32.8	504	1 HCK_MOUSE	O95m30 mecaea fasc
41	42	32.8	626	1 PRNB_MCTU	P71584 mycobacteri
42	42	32.8	675	1 BMX_HUMAN	P51813 homo sapien
43	42	32.8	693	1 PKN1_MYXA	P33277 myxococcu
44	42	32.8	766	1 GAP1_SCHNO	P33277 schizosacch
45	42	32.8	966	1 VIA_BBYV	O00020 broad bean
46	42	32.8	1100	1 JAK3_RAT	O63272 rattus norv

ALIGNMENTS

Result No.	Score	Query Match	Length	ID	Description
1	126	98.4	98	1 PFPA_ENTMH	P34095 entamoeba h
2	125	97.9	97	1 PFPA_ENTMH	O07831 entamoeba h
3	125	97.9	97	1 PFPA_ENTMH	O24824 entamoeba h
4	76	59.4	101	1 PFPA_ENTMH	O97843 syncytiotrophoblast
5	54	42.2	521	1 SPKA_SYN3	P03299 varicella-z
6	53	41.4	835	1 OBP_YZVD	P16277 mus musculus
7	50	39.1	498	1 BLK_MOUSE	P42683 gallus gall
8	49	38.3	507	1 LCK_CHICK	O92949 hellicobacte
9	48.5	37.9	559	1 PRIM_HELLO	P56004 icellulid h
10	47.5	37.1	559	1 PRIM_HELLO	O58130 icellulid h
11	46	35.9	535	1 GGP_METHA	P23561 saccharomy
12	46	35.9	535	1 GGP_METHA	P23561 saccharomy
13	46	35.9	535	1 GGP_METHA	P23561 saccharomy
14	46	35.9	535	1 GGP_METHA	P23561 saccharomy
15	45.5	35.5	660	1 EPAL_HUMAN	P73207 mycoplasma
16	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
17	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
18	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
19	45	35.2	445	1 Y390_MYCN	P54956 bacillus su
20	44.5	34.8	441	1 Y390_MYCN	P54956 bacillus su
21	44.5	34.8	441	1 Y390_MYCN	P54956 bacillus su
22	44	34.4	261	1 YGSI_YEAST	O87411 bruceella me
23	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
24	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
25	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
26	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
27	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
28	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
29	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
30	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
31	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
32	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me
33	44	34.4	379	1 YGSI_YEAST	O87411 bruceella me


```
DR      PIR: G27344; WABE51.  
DR      InterPro: IPR003450; herpes_orf_Lbp.  
DR      RefSeq: PF02399; herpes_orf_Lbp.  
DR      PubMed: 6718961; DNA binding, AP (promotional).  
FT      NcBIID: 6718961  
SO      SEQUENCE     835 AA: 94374 MW: 471P877ACB366BF CRC64:  
  
Query Match          41.4% Score 53: DB 1: Length 835;  
Query Local Similarity 43.5%; Pred. No. 3;  
Matches    10: Conservative   5: Mismatches      8: Indels       0: Gaps        0:  
  
OY      1 GFATLCTKYLKDPCDITGLDTE 23  
|| || | : | ||| |:::||  
Db      22S QFGSRCTTILDMQCIDTLVAVIK 247  
  
RESULT_7  
BLK_MOUSE BLK_MOUSE STANDARD: PROT: 498 AA.  
AD PI6277  
DT 01-AUG-1990 (Ref. 15, Created)  
DT 01-NOV-1995 (Ref. 32, Last sequence update)  
DT 15-JUN-2002 (Ref. 41, Last annotation update)  
DE Tyrosine-protein kinase Blk [EC 2.7.1.112] (B lymphocyte kinase) [655-Blk].  
GN B Lymphocyte Kinase (Blk)  
GC Eukaryota Metazoa Chordata Craniata Vertebrata Euteleostomi: Muricora Actinoptera Clupeiformes Clupeidae Scombroidei: Scombrinae: Mus. NCBJ_FaxID=10090.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=B-cell;  
KC MEDLINE=90117147; PubMed=2104338;  
RT Specific expression of a tyrosine kinase gene, blk, in B lymphoid cells".  
RT Science 247:333-336(1990).  
RN [2]  
RP STRUCTURE BY NMR OF SH2 DOMAIN.  
RX MEDLINE=96248189; PubMed=6539560;  
RY Mueller M.J., Deltling B., Pryor K., Mueller L., Farmer B.T. II,  
RZ Amino acid structural solution structure of the SH2 domain from p53-like kinase".  
CC Biochemistry 35:6201-6211(1996).  
CC -1- FUNCTION: BLK MAY FUNCTION IN A SIGNAL TRANSDUCTION PATHWAY THAT IS RESTRICTED TO B LYMPHOID CELLS.  
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.  
CC -1- SUBSTRATE: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC  
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Biomedicine and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way altered and this statement is not removed. Change by email for commercial applications or send an email to license@ebi.ac.uk (see http://www.isb.ac.uk/announce/cc)  
CC CC  
CC EMBL: M30903; AAA0453.1; ?  
DR PIR: A40092; AAA0092.  
DR PDB: 1BJL; 12-MAR-97.  
DR PDB: 1BJL; 12-MAR-97.  
DR MOD: MC1:08180; 515. Elk-pkinase.  
DR DR InterPro: IPR000980; SH2.  
DR InterPro: IPR001452; Tyr_pKinase.  
DR Pfam: PF00017; SH2; 1.  
DR Pfam: PF00018; SH3; 1.
```

	DR	Pfam: PF00069; Phkinase_1.
	DR	PRINTS: PR00401; SH2DOMAIN.
	DR	PROSITE: PS00452; SH2KINASE.
	DR	ProDom: PD000001; SH2KINASE.
	DR	ProDom: PD000066; SH3_1 Inasec. 1.
	DR	ProDom: PD000093; SH2_1.
	DR	SMART: SMO0232; SH2_1.
	DR	SMART: SMO0326; SH3_1.
	DR	SMART: SMO0219; Tyrcg. 1.
	DR	PROSITE: PS00109; PROTEIN_KINASE_ATP_1.
	DR	PROSITE: PS00109; PROTEIN_KINASE_DOM_1.
	DR	PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
	DR	PROSITE: PS50002; SH2_1.
	KW	Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;
	KM	Mystate; SH2 domain; SH3 domain; 3D-structure.
	FT	INTRMET 0
	FT	LIPID 0
	FT	DOMAIN 5
	FT	NR_BIND 117
	FT	DOMAIN 213
	FT	SH2_1
	FT	ATP (BY SIMILARITY).
	FT	NR_BIND 234
	FT	ATP (BY SIMILARITY).
	FT	NR_BIND 240
	FT	ATP (BY SIMILARITY).
	FT	BINDING 262
	FT	ACT_SITE 353
	FT	MOD_RES 382
	FO	SEQUENCE 496 AA; 36513 MW; R439D/R079F/D0577 C/GC64;
		Phosphorylation (AUTO-) (BY SIMILARITY)
		Query Match Score 50: DB 1: Length 498:
		Best Local Similarity 53.3%: Pred. No. 5.1:
		Matches 8: Conservative 4: Mismatches 3: Indels 0: Gaps
Oy		5 TLTKVDFGFIKL 19 111 1: 111: : :
Dy		364 TLCKADPELRLII 378
		RESULT 8
	ID	LCK_CHICK STANDARD: PRT: 507 AA.
AC	P42683;	
DT	01-NOV-1995 (Rel. 32, Created)	
DE	01-NOV-1995 (Rel. 32, Last sequence update)	
DE	15-JUN-2003 (Rel. 40, Last sequence update)	
DE	Proto-oncogene tyrosine-protein kinase Lck (EC 2.7.1.113) (Protein-	
DE	tyrosine kinase C-TkI).	
GN	LCK.	
OS	Gallus gallus (Chicken).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Archosauria; Mesarchosauria; Neognathae; Galliformes; Phasianidae; Phasianinae;	
ON	Nucleus	
OR	NCBI_TaxId=9031;	
RX	SEQUENCE OF 1-88 FROM N.A.	
RX	MEDLINE=92186854; PubMed=1545804;	
RA	Chow L., Ratcliffe M., Veillette A.;	
RA	"tkl is the avian homolog of the mammalian lck tyrosine protein	
RT	kinase gene".	
RT	Mol. Cell. Biol. 12:1226-1233(1992).	
RP	SEQUENCE OF 46-507 FROM N.A.	
RX	MEDLINE=8097370; PubMed=831053;	
RA	Sterbergh K., Mullins J.D., Bruck C., Rubenstein-Waligmann H.;	
RA	"Additional member of the protein-tyrosine kinase family: the src-	
RT	and lck-related protooncogene c-tkl.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 84:8778-8782(1987).	
CC	- FUNCTION: MAY PARTICIPATE IN ANTIGEN-INDUCED T-CELL ACTIVATION.	
CC	- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein	
CC	- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER	
CC	CD4 OR CD8.	
CC	- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC	
CC	SUBFAMILY.	

CX	OligoHelicobacter.
OC	NCRH_TaxID=69396;}
RN	SEQUENCE FROM N.A.
RP	MEDLINE=9120577; PubMed=9923682;
RA	Singh A., Ling L.S., Sultani M.T.,
RR	Gibson R., Werberg D., Mills S.D., Jang G.F.,
RS	Trost T.J.;
RT	"Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori";
RL	Nature 371(7-18)(1995)
RU	- RNA PRIMERS FOR THE CAZAKZI PROMOTORS ON BOMBYCELL STRAINS AT REPLICATOR FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
CC	- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
CC	- SUBUNIT: MONOMER (BY SIMILARITY).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at Geneva. The protein has been described previously by others and its use by non-profit institutions without charge are no restrictions on its modified and this statement is not removed usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/announce/ or send an email to license@isdb.slb.ch). -----
CC	EMBL: AF001441; MAD05592.1; -
DG	HSPC_OXA00020
DR	Lipid-protein complex: DNAPrimase, topPrim.
DR	Interpro: IPRO02694; ZnF-CMC2.
DR	Pfam: PF01751; Topprim_1.
DR	Pfam: PFO1807; znf-CMC2_1.
DR	ProDom: PD002968; znf_CMC2_1.
DR	SMART: SM00493; TOPPRIM_1.
DR	SMART: SM00400; ZNF CMC2_1.
DR	TransPaser: DNA primase; DNA-directed RNA polymerase; Primosome; 2N.Pinger; 3' end; 5'
FT	Zn-Finger; 37 aa; 61
KM	2N.Finger; 37 aa; 61
SO	SEQUENCE 559 AA: 63815 MW: 78139AFACB9A4F87 CRC64:
QY	Query Match 37.9%; Score 48.5; DB 1; Length 559; Best local similarity 4.8%; Freq: No. 9.7; Matches 1: Conservative 3; Mismatches 2; Indels 7; Gaps 1. : 6 LGKYLDCRDI-----DKIOL 21 Db 149 LCNKRYGIENGLMKKKOKIEL 171 :: ::
RESULT 10	
PRIM_HELICOBACTER_PYLORI_HELPY	STANDARD: PROT. 559 AA.
AD PS506.1	
DT 01-NOV-1997 (Ref. 35, Created)	
DT 01-NOV-1997 (Ref. 35, Last sequence update)	
DT 16-OCT-2001 (Ref. 40, Last annotation update)	
DN DNA primase [EC 2.7.7.-].	
DE DMS or HPOD12.	
GN Helicobacter pylori (Campylobacter pylori);	
GS Helicobacter Proteobacteria; epsilon subdivision; helicobacter group;	
NCBI_TaxID=210;	
(1)	
SEQUENCE FROM N.A.	
STRAIN=J6595 / ATCC 700392;	
MEDLINE=G739467; PubMed=9352185;	
Toim J.-E., White O., Kerttunen A.R., Clayton R.A., Sutton G.G.,	
Lehtinen M., Niemelä V., Teichgraber U.K., Kleitman R., Gill S., Dougherty B.A.,	
Loftus Jr., Richardson D., Dodson R., Khalak H.G., Glodde A.,	
McKenney K., Fitzgerald L.M., Lee N., Adams M., Hickey E.K.,	
Berg D.E., Corcoran J.D., Utterback T.R., Peterson J.D., Kelley J.M.,	

DR EMBL: M75166; AAA88136.1; -.
DR PIR: G36789; G36789.
KW Hypothetical protein.
SQ SEQUENCE 290 AA: 347.73 MW: 21C26260985ATASE CRC64;
Query Match 56 7%; Score 47; DB 1: Length 290;
Best Local Similarity: 75.0% (Protein No. 8.3);
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 GFATLCTKYVD 12
Db 217 GFITTCRKEID 228
||| ||| |||

RESULT 12
GCP_MER1A
1D GCP_MER1A STANDARD: PROT: 535 AA.
AC OS85310:
DT 15-JUL-1998 (Rel. 37, Created)
DT 15-MAR-2003 (Rel. 37, Last sequence update)
DT 15-MAR-2003 (Rel. 41), Last annotation update)
DE Putative class II polyketide chain amide synthetase (EC 3.4.24.57)
DE (glycylprotease).
GN K01130.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus;
RN NCBI_TaxId=2190;
RA SOURCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96313999, PubMed=688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischman R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Getch C.J.D.,
RA Overbeek A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Beeson
RA Kierstead T.R., Kirkness E.F., Melnick S.L., Peterson J.E., Smith
RA Cotton M.D., Roberts K.M., Harte M.N., Kaine B.P., Morodovsky M.,
RA Klek H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RL *complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii. Science 273:1058-1073(1996).
RL science 273:1058-1073(1996).

-1- FUNCTION: COULD BE A METALLOPROTEINASE.
CC -1- INTERACTS WITH 31 OTHER ASPT32 bond
CC -1- INTERACTS WITH 31 OTHER ASPT32 bond
CC unglycosylated proteins, desialylated glycoproteins or
CC glycoproteins that are only N-glycosylated.

-1- COPACITOR: ZINC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22.
CC -1- SIMILARITY: IN THE C-TERMINAL, SOME SIMILARITY WITH CONSERVED
CC REGION OF CATALYTIC DOMAIN OF PROTEIN KINASES.

CC THIS SWISS-PROT entry is copyright 1. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: U67555; AA099132.1; .
DR MEROPS: M22_UMP; .
DR TIGR: M01130; .
DR InterPro: IPR000719; Euk_pHase.
DR InterPro: IPR000905; Peptidase_M22.
DR InterPro: IPR001245; Tyr_pHase.
DR Pfam: PF00814; Peptidase_M22.1.
DR PRINTS: PR00769; OSALOPEASE.
DR PROSITE: PS00289; Peptidase_M22.1.
DR Problem: P0002367; Peptidase_M22.1.
DR TRGFAMS: TIC801029; gcp: 1 -M22.1.
DR PROSITE: PS01016; GLYCOPROTEINASE.1

DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN.1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; UNKNOWN.1.
 DR CYTOPLASMIC PROTEIN; HYDROLASE; METALLOPROTEINASE; ZINC;
 KM CYTOPLASMIC PROTEIN; HYDROLASE; METALLOPROTEINASE; ZINC;
 FT DOMAIN 1 312 METALLOPEPTIDASE.
 FT DOMAIN 333 535 PROTEIN KINASE LINK.
 FT METAL 105 106 ZINC (POTENTIAL).
 FT METAL 110 110 ZINC (POTENTIAL).
 SO SEQUENCE 535 AA; 60570 MW; E1BEEFACEB55999F CRC64;
 Query Match 35.9%; Score 46; DB 1; Length 535;
 Best local similarity 50.0%; Pseq No. 22;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 Oy 10 VIDFIDGKIDLTLDLEK 25
 Db 465 IDFGGLKISMDLDR 480
 RESULT 13
 ID STILL_YEAST STANDARD; PRT: 738 AA.
 AC P23561;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 06-OCT-2001 (Rel. 40, Last annotation update)
 DT 06-OCT-2001 (Rel. 40, Last annotation update)
 GN STELL OR YLR162M OR I6039.10
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 NC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NX NCBI_TaxID=4932;
 RM 11 DUNCE OF 23-738 FROM N.A.
 RM MEDLINE=9115076; PubMed=4276621;
 RA Rhodes N., Connell I., Errede B.;
 RT STELL is a protein kinase required for cell-type-specific
 transcription and signal transduction in yeast.*;
 RL Genes Dev. 4:1862-1874(1990).
 ID DUNCE FROM N.A.
 AC STRAINS=286; 1 Apr92;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Johnston A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hallier L., Jier W.,
 RA Johnson D., Johnston L., Langston Y., Lattelle P., Le T.,
 RA Marids E., Meneses S., Miller N., Nham M., Pauley A., Peluso D.,
 RA Ritten L., Riles L., Tatch A., Trevaaks E., Vignati D.,
 RA Ritten L., Woltemen P., Vaidin M., Wilson R., Waterston R.;
 RA Smithed (CAN-1995) to the EMBL/GenBank/DBD databases.
 RM 13
 RM POSSIBLE FUNCTION:
 RX MEDLINE=9233193; PubMed=1628833;
 RA Cairns B.R., Ramer S.W., Kornberg K.D.;
 RT Order of action of components in the yeast pheromone response
 pathway revealed with a dominant allele of the STE11 kinase and the
 STELL protein kinase.
 RL Genes Dev. 6:1055-1100(1992).
 ID SPECIFIC TRANSCRIPTION AND SIGNAL TRANSDUCTION IN YEAST. IT
 IS THOUGHT THAT IT PHOSPHORYLATES THE STE11 KINASE WHICH
 ITSELF, PHOSPHORYLATES THE FUS3 AND OR KSS1 KINASES.
 CC -1. PATHWAY: RESPONSE TO PHEROMONE-INDUCED SIGNAL.
 CC -1. SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1. SIMILARITY: KINASE KINASE SUBFAMILY.
 CC -1. SIMILARITY: CONTROLLED BY STE11.
 CC THIS SWISS-PROT entry is copyright, it is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 distorted and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC EMBL: X53431; CAAT522.1; -;
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; UNKNOWN.1.
 DR PROSITE: PS500109; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50105; SAM_DOMAIN.1.
 DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
 KM TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
 FT DOMAIN 436 733 SAM
 FT DOMAIN 442 450 ATP (BY SIMILARITY).
 FT BINDING 465 465 ATP (BY SIMILARITY).
 FT ACT SITE 600 BY SIMILARITY.
 SO SEQUENCE 738 AA; 83217 MW; A5006900B346A1 CRC64;
 Query Match 35.9%; Score 46; DB 1; Length 738;
 Best local similarity 42.1%; Pseq No. 31;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 Oy 7 CTKVDLFDGKIDLTLDLEK 25
 Db 613 CVKIDFDGKISKSLPLKK 631
 RESULT 14
 ID EPAL_HUMAN STANDARD; PRT: 976 AA.
 AC P21709; 015405;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 11-OCT-2001 (Rel. 40, Last annotation update)
 GN Ephrin type-A receptor 1 precursor (Ec 2.7.1.112) (tyrosine-protein
 kinase receptor Eph).
 GN EPHAL OR EPHR1 OR EPHR OR EPH.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RM 11
 RM SEQUENCE FROM N.A.
 RX MEDLINE=88070650; PubMed=2825356;
 RA Hirai H., Maru Y., Hagihara K., Mitsuhashi T., Takaku F.;
 RT A novel putative tyrosine kinase receptor encoded by the eph gene.*;
 RL Science 258:1171-1172(1997).
 ID DUNCE FROM N.A.
 AC MEDLINE=92929440; PubMed=10369740;
 RA Owehalmipur D., Kelley M.J.;
 RT Genomic structure of the EPAL receptor tyrosine kinase gene.*;
 RL Mol. Cell. Probes 13:169-173(1999).
 RN (3)
 RN SEQUENCE OF 286-976 FROM N.A.
 RM TISSUE=placenta;
 AC TISSUE=placenta;
 RA SUPPLEMENTARY DATA (MAY-1994) to the EMBL/GenBank/DBD databases.
 CC -1. FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS WITH
 CC A LOW AFFINITY TO EPHRIN-A1.
 CC -1. CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1. SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1. TISSUE SPECIFICITY: OVEREXPRESSED IN SEVERAL CARCINOMAS.

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OK protein - protein search, using sw model

Run on: May 9, 2003, 16:10:40 : Search time 18.5294 seconds
(without alignments)
300.240 Million cell updates/sec

Title: us-09-851-422a-8
Perfect score: 128
Sequence: 1 GFLATCTKVLDFGIDKLQIQLIEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-unclassified:*
- 14: sp-unclassified:*
- 15: sp-virus:*
- 16: sp-bacteriophage:*
- 17: sp-archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	88	68.8	96	099855	099855
2	69	53.9	101	099854	099854
3	63	49.2	87	015591	015591
4	54	42.2	421	10 094909	094909
5	54	42.2	459	10 094300	094300
6	53	41.4	835	12 093397	093397
7	52	40.6	366	17 026743	026743
8	51	39.8	520	16 089752	089752
9	51	39.8	213	16 094314	094314
10	50	39.1	233	5 094112	094112
11	50	39.1	816	12 094199	094199
12	49	38.3	309	3 008743	008743
13	49	38.3	629	10 094304	094304
14	48	37.5	371	5 023526	023526
15	48	37.5	371	5 023526	023526

17	48	37.5	425	16 0998P5	0998P5
18	48	37.5	443	10 0943D1	0943D1
19	48	37.5	481	16 072727	072727
20	48	37.5	628	10 094331	094331
21	48	37.5	628	10 094331	094331
22	47.5	37.1	1885	10 0943C6	0943C6
23	47.5	37.1	1885	10 0943C6	0943C6
24	47	36.7	131	16 092M33	092M33
25	47	36.7	183	16 092UM3	092UM3
26	47	36.7	317	10 094M21	094M21
27	47	36.7	372	5 094M21	094M21
28	47	36.7	372	5 094M21	094M21
29	47	36.7	372	5 094M21	094M21
30	47	36.7	372	5 094M21	094M21
31	47	36.7	372	5 094M21	094M21
32	47	36.7	372	5 094M21	094M21
33	47	36.7	372	5 094M21	094M21
34	47	36.7	537	10 0943D6	0943D6
35	47	36.7	618	10 0943D9	0943D9
36	47	36.7	806	10 0943D5	0943D5
37	46.5	36.3	287	5 0947C4	0947C4
38	46.5	36.3	258	5 087M44	087M44
39	46.5	36.3	379	2 092EX1	092EX1
40	46.5	36.3	859	16 0946P4	0946P4
41	46	35.9	131	16 025024	025024
42	46	35.9	131	16 025024	025024
43	46	35.9	131	16 025024	025024
44	46	35.9	131	16 025024	025024
45	46	35.9	131	16 025024	025024

ALIGNMENTS

Result ID	Query Match	Length DB	ID	Description
1	88	96	099855	099855
2	69	101	099854	099854
3	63	87	015591	015591
4	54	421	10 094909	094909
5	54	459	10 094300	094300
6	53	835	12 093397	093397
7	52	366	17 026743	026743
8	51	520	16 089752	089752
9	51	213	16 094314	094314
10	50	233	5 094112	094112
11	50	816	12 094199	094199
12	49	309	3 008743	008743
13	49	629	10 094304	094304
14	48	371	5 023526	023526
15	48	371	5 023526	023526

0908G4 ID 0908G4 PRELIMINARY: PRT: 101 AA.
 AC 0908G4-2000 (TREMblrel. 13. Created)
 DT 01-JUN-1998 (TREMblrel. 05. Last annotation update)
 DT 01-JUN-2002 (TREMblrel. 21. Last annotation update)
 DE Pore-forming protein Isoform C precursor.
 DN DP-C.
 OS Entamoeba dispar.
 CC Eukaryota; Entamoebidae; Entamoeba.
 NC NCBI_TaxID=45681;
 NM [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SAW 142;
 RX MEDLINE=99449604; PubMed=10518795;
 RA Nickel R., Ott C., Dandekar T., Leipe M.;
 RT "Pore-forming peptides of entamoeba dispar similarity and divergence
 RT of amoebae in structure, expression and activity."
 DR EMBL: AF082528; AF041961.
 DR InterPro: IPR000004; SAPB.
 DR ProDom: PD001732; SAPB: 1.
 DR SMART: SM00118; SAPB: 1.
 KW SIGNAL.
 KW SIGNAL.
 FT CHAIN 1 25 10877 MW: 7699kDa;SD08/6EB CRC4;
 SC SEQUENCE 101 AA: 10877 MW: 7699kDa;SD08/6EB CRC4;
 50
 Query Match 53.9%; Score 69; DB 5; Length 101;
 Best Local Similarity 54.5%; Pred. No. 0.0053;
 Matches 12; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
 0Y 1 GFIATCTVLDGFDKLIQLI 22
 DB 64 GVIECVSNIVSGIDKLEKI 85
 RESULT 3
 ID 015591 PRELIMINARY: PRT: 87 AA.
 AC 015591-2001 (TREMblrel. 05. Created)
 DT 01-JAN-1998 (TREMblrel. 05. Last annotation update)
 DT 01-JAN-1998 (TREMblrel. 20. Last annotation update)
 DE AmoebaPorte C homologue (Fragment).
 DN AmoebaPorte C homologue (Fragment).
 OS Entamoeba histolytica.
 CC Eukaryota; Entamoebidae; Entamoeba.
 NC NCBI_TaxID=5759;
 NM [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HM-1;
 RX MEDLINE=97396144; PubMed=9245698;
 RA Tanaka T., Tanaka M., Mitsui Y.;
 RT "Analysis of expressed sequence tags (ESTs) of the parasitic protozoa
 RT Entamoeba histolytica."
 DR EMBL: AB002741; BA219861.
 DR InterPro: IPR000004; SAPB.
 DR SMART: SM00118; SAPB: 1.
 FT NON_TER 1 87
 FT NON_TER 87 87
 SC SEQUENCE 87 AA: 9131 MW: FDDIC1BOFAABBB6E CRC4;
 50
 Query Match 49.2%; Score 63; DB 5; Length 87;
 Best Local Similarity 54.5%; Pred. No. 0.033;
 Matches 12; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 0Y 1 GFIATCTVLDGFDKLIQLI 22
 DB 59 GIVATCTVLDGFDKLIQLI 80
 RESULT 4

09409 ID 09409 PRELIMINARY: PRT: 421 AA.
 AC 09409-1998 (TREMblrel. 06. Created)
 DT 01-JUN-1998 (TREMblrel. 05. Last annotation update)
 DT 01-JUN-1998 (TREMblrel. 21. Last annotation update)
 DE Protein kinase - like protein (Protein kinase-like protein).
 DN P1KCS_120 OR ATG18950.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 CC NCBI_TaxID=3702.
 NM [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Pohl T., Weizsenger T., Bancroft I., Mewes H.W., Mayer K.,
 RA Schuller C.;
 RI Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
 SC SEQUENCE FROM N.A.
 RA Pohl T., Weizsenger T., Mewes H.W., Lemcke K., Mayer K.F.X.;
 RI Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RI Submitted (Mar-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL161543; CAB78897.1.
 DR InterPro: IPR002110; ANK.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR004040; STY_Pkinase.
 DR Pfam: PF00023; ANK: 3.
 DR Pfam: PF00065; Pkinase: 1.
 DR ProDom: PD00065; Pkinase: 1.
 DR SMART: SM00248; ANK: 1.
 DR SMART: SM00221; STYK: 1.
 DR PROSITE: PS50088; ANK_REPEAT: 2.
 DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
 DR ANK_repeat: ATP-binding kinase; Repeat; Transferrase.
 SC SEQUENCE 421 AA: 40173 MW: 6600787.868658kDa CRC4;
 50
 Query Match 42.8%; Score 54; DB 10; Length 421;
 Best Local Similarity 58.8%; Pred. No. 4.6;
 Matches 10; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 0Y 9 KVDGFDKLIQLIQLEK 25
 DB 258 KVDGFDKLIQLEK 274
 RESULT 5
 ID 093230 PRELIMINARY: PRT: 459 AA.
 AC 093230-2001 (TREMblrel. 15. Created)
 DT 01-DEC-2001 (TREMblrel. 15. Last annotation update)
 DT 01-JUN-2002 (TREMblrel. 21. Last annotation update)
 DE Arabidopsis thaliana (Mouse-ear cress).
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
 CC NCBI_TaxID=3702.
 NM [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T.M., Kamaya A., Karlin-Neumann G.,
 RA Kawai J., Lam J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satoh M., Seki M., Sotomiyama A., Tang C.C., Tortum M., Tanaka K.,
 RA Yamamura T., Yu G., Yu S., Yoshizaki K., Davis R.W., Theologis A.,

```

RA Ecker J.R.:
RA "Ariadopsis is cDNA clones"
RA Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
RL ID: 256211
RL Accession: AF060211
RL InterPro: IPR002110: RNK
DR InterPro: IPR000719: Euk_Pkinase
DR Pfam: PF00023: ank; 3
DR Pfam: PF00069: pkinase; 1
DR ProDom: PD000001: Euk_pkinase; 1
DR PROSITE: PSS0088: ANK_REPEAT; 2
DR PROSITE: PSS0297: ANK_REPEAT_REGION; 1
DR PROSITE: PSS0133: PROTEIN_KINASE_DOM; 1
RA NCBI_taxID: 187420: Methanobacteriales;
SO SEQUENCE 459 AA: 52617 MW: 6170458552252R02 CMG64:

Query Match 42.2%; Score 54; DB 10; Length 459;
Best local similarity 58.8%; Pval. No. 5;
Matches 10: Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 9 KYLDGKIXLQILDEK 25
DB 296 KVADFGVSLVTVKEDK 312

RESULT 6
OY 09J3N7 PRELIMINARY: PRT: 835 AA.
AC 09J3N7
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE ORF51.
OS Human herpesvirus 3
OC Alphaherpesvirinae
OC Herpesviridae
OC Nucleocapsid
OC Nucleocapsid
OC NCL_TaxID:10335;
RN [1]
RA SEQUENCE FROM N.A.
RA Argay T., Cohen J.I., Klutch M., Lokstrom K., Yoshikawa T., Asano Y.,
RA Krause P.R.:
RA "Nucleotide sequences that distinguish OKa vaccine from parental OKa
RA and other varicella-zoster virus isolates."
RA J. Virol. 73:303-307, 1999.
RA EMBL: AF206304: AF6153; 1
RA InterPro: IPR003450: Herpes_crl_bp.
DR Pfam: PF02399: Herpes_crl_bp; 1.
SO SEQUENCE 835 AA: 94383 MW: B0368877ACEA9FFB CMG64:

Query Match 41.4%; Score 53; DB 12; Length 835;
Best local similarity 44.8%; Pval. No. 13;
Matches 10: Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 GFATICTVVDGIDKIQITE 23
DB 225 GFSRRTCTILDMGIDTLVAVIK 247

RESULT 7
OY 026243 PRELIMINARY: PRT: 396 AA.
AC 026243
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Coded protein.
OS Methanobacterium thermoautotrophicum.
OC Archaea: Euryarchaeota: Methanobacteriales;
OC Methanobacteriales: Methanobacteriales;
OC NCBI_taxID:187420:
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN:DELTA H;

```

```

RX MEDLINE:98037514; PubMed:9371463;
RX Smith D.R., Doucoule-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RX Alredette D., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RX Harrison D., Hwang L., Keagle P., Lamm W., Pochter B., Qiu D.,
RX Jentoft A., Kozare R., Ning Y., Pietrowski S., Gibson R.,
RX Jentoft A., Canessa C., Hwang L., Keagle P., Lamm W., Pochter B.,
RX McQuigill S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
RX Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.:
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RA delta: Functional analysis and comparative genomics."
RL J. Bacteriol. 179:7133-7155(1997).
DR EMBL: AE008003: AAB84646.1;
DR InterPro: IPR003741: Duf182.
DR Pfam: PF02569: Duf182; 1.
DR PROSITE: PSS0037: ferf; 2.
DR Iron-sulfur: Complete proteome.
SO SEQUENCE 396 AA: 42840 MW: 5800487232499B CMG64:

Query Match 40.6%; Score 52; DB 17; Length 396;
Best local similarity 40.9%; Pval. No. 8.9;
Matches 9: Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 3 IATICTRVLDGIDKIQILIED 24
DB 207 LSLMDTHIVGVCKLDPOLIED 228

RESULT 8
OY 08TV25 PRELIMINARY: PRT: 491 AA.
AC 08TV25
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Anthic acid transposase.
OS Methanopyrus kandleri.
OC Archaea: Euryarchaeota: Methanopyri: Methanopyriales; Methanopyraceae;
OC Methanopyrus.
OC NCBI_taxID:3320:
RA SEQUENCE FROM N.A.
RA STRAIN:AV19 / DSM 6324 / JCM 9639;
RX MEDLINE:21927647; PubMed:11930012;
RA Slesarev A.I., Meshvaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbakina O.V., Shakhova V.V., Belova G.I., Arsyand L.,
RA Natarle D.A., Rogozin I.B., Tolstov R.L., Wolf Y.I., Stetter K.O.,
RA Nayak A.G., Kozman E.V., Kozman S.A.:
RA "Genomic and proteomic analysis of the Methanopyrus kandleri AV19
RA and monopyr of archaeal methanopyrus."
RT Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL: AE010447: AAM02789.1;
KW Complete proteome.
SO SEQUENCE 491 AA: 52944 MW: 0A087ED3E2981B9B CMG64:

Query Match 39.8%; Score 51; DB 17; Length 491;
Best local similarity 40.0%; Pval. No. 10;
Matches 10: Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 1 GFATICTVVDGIDKIQILDEK 25
DB 14 GFLSTICTVIGGVNVLQCMIOVK 38

RESULT 9
OY 08YNP6 PRELIMINARY: PRT: 520 AA.
AC 08YNP6
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

```

DE Serine/threonine kinase.
 ON ALL4518.
 OS Anabaena sp. (strain PCC 7120).
 OC Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NX NCBI_TaxID=103690.
 RN 111
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Molk C.P., Kuriitz T., Sasamoto S.,
 RA Matsumoto A., Iriuguchi M., Ishikawa A., Kawashima K., Kikuta T.,
 RA Minamide Y., Kohira M., Matsuno M., Katsuhara A., Muraki A.,
 RA Yashida M., Tabata S., Sugimoto M., Takazawa M., Tamada M.,
 RA "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RT DNA Res. 8:205-213(2001).
 DR EMBL: AB003596; BAB76217.1.
 DR InterPro: IPR002200; Euk-kinase.
 DR InterPro: IPR002200; Euk-kinase.
 DR Pfam: PF00609; Pkinase.1.
 DR ProDom: P0000001; Euk_Pkinase.1.
 DR SMART: SMO0220; S_TKc.1.
 DR SMART: SMO0219; TYRCK.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP_UNKNOWN_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00109; PROTEIN_KINASE_ST.1.
 KM Kinase. Complete proteome
 SO SEQUENCE 520 AA; 58905 MW; 26FE259B80ACF49 CRC64;

Query Match 39.8%; Score 51; DB 16; Length 520;
 Host Local Similarity 64.7%; Pred. No. 17;

Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Oy 9 KYLDGIDKLTQLIEDK 25
 |||||:||||:||||:|
 Db 168 KYLDGIDKLTQLIEDK 184

RESULT 10
 ON 09A314 PRELIMINARY; PRT: 213 AA.
 AC 09A314.
 DT 01-JUN-2001 (TRENBLrel. 17, Created)
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein CC3392.
 DB CC3392
 OS Caulobacter crescentus
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 CC Caulobacter.
 NX NCBI_TaxID=155892;
 RN 111
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 10699 / CH15.
 RA MEDLINE=21359213; PubMed=11259647;
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N.,addock J.R.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gilman M.L., Haft D.H.,
 RA Kolonchak J.F., Salt J., Craven M.B., Khout H., Shetty J., Berry K.,
 RA Kutchback T., Tan K., Wolf A., Vannakaveh J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RA "Complete genome sequence of Caulobacter crescentus";
 RT Science 287:2185-2195(2000).
 DR EMBL: AE005999; AAK5554.1.
 DR EMBL: AE005999; AAK5554.1.
 DR TIGR: CC3392;
 KM Hypothetical protein: Complete proteome.
 SO SEQUENCE 213 AA; 22524 MW; 609F9FA1ADMAZE CRC64;

Query Match 39.1%; Score 50; DB 16; Length 213;
 Best Local Similarity 40.0%; Pred. No. 9.7;
 Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Oy 1 GFATLCTKVLDFGIDKLTQLIEDK 25
 |||||:||||:||||:|
 Db 177 GFSASTTKRSDFGVSKYTPWVGE 201

RESULT 11
 ON 09V112 PRELIMINARY; PRT: 233 AA.
 AC 09V112.
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE CG34400 protein.
 DB CG34400
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 NX NCBI_TaxID=7227.
 RN 111
 RP SEQUENCE FROM N.A.
 RA STRAIN=BERKELEY.
 RA MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga P.G., Scherer S.E., Li P.M., Hoskins R.A., Gale R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Breen J., Cawley S., Daly M.J., DeLuca C., Drenth J., Eick J.,
 RA Brandon R.C., Rogers Y.-H.C., Blazed R.G., Champe M., Pfeiffer B.D.,
 RA Man R.H., Doyle C., Baxter E.G., Holt C., Nelson C.R., Miklos G.L.C.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolintaru S.,
 RA Botcheva D., Botchan M.R., Bouck J., Brockstein P., Broderick P.,
 RA Brown J.M., Brown J.A., Bruner R.D., Busch J., Butler G.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Durbin K.J., Evangelista C.C., Farrar C., Fertire S., Fleischmann W.,
 RA Foster A.C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris R.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
 RA Holt R.A., Houston K.A., Hunkeler D., Hunkeler D., Hunkeler D.,
 RA Jaiswal J.M., Kaul S., Kaul S., Kaul S., Kaul S., Kaul S., Kaul S.,
 RA Kimmel B.E., Kodly C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Lin X., Mathe B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J.R., Moshrefi A.,
 RA Nelson K.L., Nguyen M., Nguyen M., Nguyen M., Nguyen M., Nguyen M.,
 RA Palazzolo M., Pittman G.S., Pon S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassman D.A., Weinstock G.M., Weissbach J.,
 RA Ye J., Yeh J., Yeh J., Yeh J., Yeh J., Yeh J., Yeh J., Yeh J.,
 RA Zhang X.H., Zhong F.N., Zhong M., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003668; AAF3937.1.
 DR EMBL: AE003668; AAF3937.1.
 DR EMBL: AE003668; AAF3937.1.
 SO SEQUENCE 233 AA; 27238 MW; 4189178B71A1714 CRC64;

Query Match 39.1%; Score 50; DB 5; Length 233;
 Best Local Similarity 47.6%; Pred. No. 11;
 Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 Oy 4 ATLCTKVLDFGIDKLTQLIED 24
 |||||:||||:||||:|
 Db 5 ATLCTKVLDFGIDKLTQLIED 25

Query Match Similarity 38.3% Score 49; DB 3 Length 309;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 ATCTKVALDFGDKL 19
|||||
DB 86 ATLCTNLILFGSLII 101

RESULT 14

ID O943D4 PRELIMINARY; PRT. 629 AA.
DT 01-DEC-2001 (TREMBL) 19. Created)
DT 01-DEC-2001 (TREMBL) 19. Last sequence update)
DT 01-JUN-2002 (TREMBL) 21. Last annotation update)
GN P043D08.19 protein.
GN P043D08.19.
OC Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Gymnosperms; Angiosperms; Liliopsida; Poales; Poaceae;
OC Brachiopsida; Poaceae; Oryza.
OX NCBI_TaxID=4530,
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa niponbare(GAS) genome DNA, chromosome 1, PAC
clone P043D08-2001) to the EMBL/GenBank/DDJJ databases."
DR EMBL: AB003250, BABF6139.1, -;
DR EMBL: AP000345; CytC_kine_bln.d.
DR InterPro: IPRO00719; Euk_Pkinase.
DR InterPro: IPRO02250; Ser_thr_pkinase.
DR Pfam: PF03578; HGM; 1.
DR Pfam: PF03578; HGM; 1.
DR ProDom: PD00060; Cytokichome_1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Transferase.

SOURCE 629 AA: Genbank: A6D7AE2753FCB3 CRC64;

Query Match 38.3% Score 49; DB 10; Length 629;
Best Local Similarity 56.2%; Pred. No. 41;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 6 LCRTKAFGIDKL 21
|||::|||::|||
DB 504 LCRTADPGRGLATICH 519

RESULT 15

ID O96X52 PRELIMINARY; PRT. 551 AA.
RX 096X52.
DT 01-DEC-2001 (TREMBL) 19. Created)
DT 01-DEC-2001 (TREMBL) 19. Last sequence update)
DT 01-JUN-2002 (TREMBL) 21. Last annotation update)
GN Phosphatase transposite.
GN P0185
OC Eukaryota; Fungi; Basidiomycota; Hymenomyces; Homobasidiomycetes;
OC Agaricales; Strophariaceae; Psilocybe.
OX NCBI_TaxID=61267;
RN [1]
RP SOURCE FROM N.A.
RC STRAIN=MZ;
RA Tasaki Y., Kamiya Y., Hara T., Joh T.;
RT "Isolation and sequencing of cDNA of genes induced by phosphate
deficiency in cultured mycelia of *Psilocybe* namoko";
submitted (Apr-2001) to the EMBL/GenBank/DDJJ databases.

```

CC      -1: SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR      EMBL: AB060641; BAB43910.1; -.
DR      DDBJ: F58049.2; AB060641.1; BLOS_Termaseg.
DR      InterPro: IPR003662; SUGAR_Transporter.
DR      PIRam: PF00083; sugat_tr_1.
DR      TIGRFamS: TIGR00887; ZAO109; 1.
DR      PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
DR      PROSITE: PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KM      Transmembrane.
SQ      SEQUENCE 551 AA: 59892 MW: 97ED9A5C7C89FC1 CRC64:
Query Match      37.9%; Score 48.5; DB 3; Length 551;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 13; Conservative 2; Mismatches 10; Indels 1; Gaps 1;
OY      1 GFPA-TICTKVLDPGIDKLIQLEEK 25
DB      402 GFPLALTLFVLIGFIDKLISTPSRK 427

```

Search completed: May 9, 2003, 16:17:39
 Job time : 21.5294 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:44 ; Search time 40.95 Seconds
(without alignments)
70,286 Malignments

Title: US-09-851-422B-1
Perfect score: 46

Sequence: 1 LVDRATCLR 9

Scoring table: BIOSIM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Maximum de seq. length: 200000000

Maximum Match 100%

Listing first 45 summaries

1. A/GATGGGTTATGTTGCGTGGGTTGGGATGGGTTAA/AA190. DAT
2. S/DSD2/GCGGdAtA/gemseq/gemseqp-emb1/AA191. DAT
3. S/DSD3/GCGdAtA/gemseq/gemseqp-emb1/AA192. DAT
4. S/DSD5/GCGdAtA/gemseq/gemseqp-emb1/AA193. DAT
5. S/DSD7/GCGdAtA/gemseq/gemseqp-emb1/AA194. DAT
6. S/DSD8/GCGdAtA/gemseq/gemseqp-emb1/AA195. DAT
7. S/DSD9/GCGdAtA/gemseq/gemseqp-emb1/AA196. DAT
8. S/DSD5/GCGdAtA/gemseq/gemseqp-emb1/AA197. DAT
9. S/DSD5/GCGdAtA/gemseq/gemseqp-emb1/AA198. DAT
10. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
11. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
12. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
13. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
14. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
15. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
16. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
17. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
18. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
19. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
20. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
21. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
22. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA199. DAT
23. S/DSD2/GCGdAtA/gemseq/gemseqp-emb1/AA200. DAT

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	181	22	AAH56008
2	46	100.0	182	22	AAH50289
3	46	100.0	185	22	AAH50289
4	46	100.0	415	14	AAH52765
5	46	100.0	415	16	AAH52765
6	46	100.0	454	16	AAH67710
7	46	100.0	456	5	AAH40128
8	46	100.0	456	5	AAH40222
9	46	100.0	451	6	AAH50311
10	46	100.0	451	6	AAH50302
11	46	100.0	451	6	AAH50019

11	46	100.0	461	11	AAW05393	Human factor IX precursor
12	46	100.0	461	19	AAW04284	Human factor IX pr
13	46	100.0	461	21	AAV97295	Human clothing fac
14	46	100.0	461	22	AAE10628	Human wild-type fac
15	46	100.0	461	22	AAW6281	Human factor IX f
16	46	100.0	461	23	AAW02152	Human factor IX f
17	46	100.0	461	23	AAW02152	Protein of human f
18	46	100.0	462	13	AAW10688	Human factor IX
19	41	89.1	415	20	AAW03203	Antio acid sequen
20	37	80.4	357	22	AAU47611	Proionibacterium
21	34	73.9	63	20	AAU11751	Human 5' EST seque
22	34	73.9	63	21	AAW00629	Human factor IX
23	34	73.9	158	12	AAU11721	Human secreted pro
24	33	71.7	96	22	AAU18643	Human factor IX
25	33	71.7	101	21	AAW46487	Novel bone marrow
26	33	71.7	138	20	AAU13395	Human gene 11-III'
27	33	71.7	138	20	AAU47425	HEV peptide 412 or
28	33	71.7	138	20	AAU47425	Proionibacterium
29	33	71.7	250	11	AAW05472	Human factor IX
30	33	71.7	250	11	AAW05472	Human factor IX
31	33	71.7	345	22	AAW63788	Human factor IX
32	33	71.7	575	22	AAW36439	Human factor IX
33	33	71.7	581	22	AAU13363	Human factor IX
34	33	71.7	821	22	AAU13265	Human factor IX
35	33	71.7	877	22	AAW63624	Human factor IX
36	33	71.7	997	12	AAW15345	Human factor IX
37	33	71.7	1174	22	AAW01235	Novel human diapo
38	33	69.6	38	14	AAW40320	ICAM-1 peptide f
39	33	69.6	38	14	AAW40320	Peptide #2014 enco
40	32	69.6	83	22	AAW29353	Peptide #2014 enco
41	32	69.6	83	22	AAW31557	Peptide #1945 enco
42	32	69.6	83	22	AAW19466	Protein #1945 enco
43	32	69.6	83	22	AAW55233	Human brain expres
44	32	69.6	83	22	AAW67720	Human bone marrow
45	32	69.6	83	22	AAW15530	Peptide #1961 enco

ALIGNMENTS

RESULT 1

ID	Peptide	AA
ABB36008	standard	181

AC ABB36008;

DT 04-FEB-2002 (first entry)
XX

Peptide #3514 encoded by human foetal liver single exon probe.

Human; foetal liver; gene expression; single exon nucleic acid probe.

US Homo sapien
XX

PN WO200157277-A2.

PD 09-AUG-2001
XX

PF 30-JAN-2001; 2001WO-US000669; XX

PR	04-FEB-2000; 2000US-0180312.
PR	26-MAY-2000; 2000US-0207456.

PR 03-AUG-2000: 2000US-0632366

27-SEP-2000: 2000US-0236359,

XX		9	7	6	5	4	3	2	1
XX		8	7	6	5	4	3	2	1

XXXXXXXXXXXXXXXXXXXX

XX

Penn SG, Hanzel DK, Chen W, Rank DR, WPI; 2001-483447/52.

CC	can increase hFIX levels over the lifespan of an individual; however, they
CC	can independently exert effects on hFIX mRNA in an age-related manner,
CC	with AEs ¹ acting to stabilize hFIX mRNA, and AEs ² acting to increase hFIX
CC	mRNA levels. Over time, AEs also directs the expression of the
CC	hFIX gene, which accounts for approximately 57% of the total hFIX
CC	(AAE508A1-3' and 5'-CAGGAG-3'. The age-related regulatory sequences of
CC	the invention, along with their homologues, variants and fragments, may
CC	be used in the construction of recombinant expression vectors for the
CC	expression of a desired gene in an age-related fashion in a host
CC	cell. Preferred target genes for expression in such age-regulatable
CC	vectors include the pro-coagulant factors IX and factor VIII, and the
CC	anti-coagulants protein C and antithrombin III), human
CC	alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
CC	luciferase. Preferred promoters for use in such age-regulatable
CC	expression vectors include the human factor IX promoter, the T7 promoter,
CC	the T3 promoter and the Sp6 promoter. The expression vectors of the
CC	invention may be used in gene therapy to provide age-related and/or
CC	age-independent regulation of such age-related conditions such as
CC	hemophilia, cardiovascular disease, diabetes, Alzheimer's disease,
CC	Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
CC	Specifically, they may be used to express factor IX antisense mRNA in the
CC	treatment of thrombotic conditions associated with the natural
CC	age-related rise in factor IX expression. Transgenic cells or animals
CC	that contain vectors of the invention are useful as models of disease
CC	and normal processes such as aging and gene expression. Fragments and
CC	homologues of age-related regulatory sequences, are useful as probes to
CC	detect, isolate or identify other such sequences in samples. The present
CC	sequence represents a fragment of hFIX.
XX	
XX	Sequence 182 AA:
XX	
XX	Query Match 100.0%; Score 46; DP 22; Length 182;
XX	Best Local Similarity 100.0%; Pident NO 0.43; Mismatches 0;
XX	Gaps 0;
YY	Matches 9; Conservative 0; Indels 0; Gaps 0;
DB	
DB	1 LVDRATCR 9
DB	
DB	97 LVDRATCR 105
RESULT 3	
AA64266	AA64266 standard: protein: 412 AA.
AC	AA64266:
AC	18-JUL-1995 (first entry)
DC	Human Factor-IX.
KM	Factor-IX: polymer conjugate; hemophilias-B; blood-clotting;
KM	coagulant; polyethylene glycol; peg.
OS	Homo sapiens.
XX	
XX	Key
FT	Domain
FT	/label= "gln.domain"
FT	/note= "gamma carboxy-glutamate-rich domain"
FT	Modified-site
FT	/label= OTHER
FT	/note= "gamma-carboxyglutamic acid"
FT	Modified-site
FT	/label= OTHER
FT	/note= "gamma-carboxyglutamic acid"
FT	Modified-site
FT	/label= OTHER
FT	/note= "gamma-carboxyglutamic acid"

[illegible]

CC	polyethylene glycol show reduced immunogenicity
XX	Sequence 412 AA:
SQ	Query Match 100.0%; Score 46; DB 16; Length 412;
DB	Biological Similarity 100.0%; P=1e-96; 0; Indels 0; Gaps 0;
Matches	9; Conservative 0; Mismatches
OY	1 LVDRAICLR 9
DB	
	327 LVDRAICLR 335
RESULT 4	
ID	AAR35761 standard; protein: 415 AA.
AC	AAR35761;
XX	AAR35761;
DT	24-SEP-1993 (first entry)
DX	
XX	Factor IX (IX).
XX	
KW	PC; protein C; IX; Factor IX; X; Factor X; PT; prothrombin; VII;
KW	Factor VIII; CT; chymotrypsinogen; SP; serine protease; binding;
KW	exosite; catalytic activity.
XX	
OS	Homo sapiens.
FT	
FT	Key Location/Qualifiers
FT	Region 1..145 /note="Factor IX light chain"
FT	Region 146..180 /note="Factor IX activation"
FT	Region 181..415 /note="Factor IX heavy chain"
FT	Peptide 181..409 /note="exosite 1"
FT	Peptide 316..336 /note="exosite 2"
FT	Peptide 316..330
FT	/note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 321..330
FT	/note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 393..409 /note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 400..409 /note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 400..414 /note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 400..414 /note=".pref. PC polypeptide; claim 2, page 136"
FT	Peptide 271..292
FT	/note="claim 6, page 137-138 describes an antibody
FT	that reacts with factor IX fragments
FT	316-330, 321-330, 393-409, 400-414
FT	400-414 but not with fragment 271-292"
PX	
XX	M03030804-A.
PX	
PD	27-MAY-1993.
PD	
PD	18-NOV-1992; 92MO-US10242.
XX	
PR	18-NOV-1991; 91US-0793989.
XX	
PA	(SCRI) SCAIPPS RES INST.
XX	
PI	Griffin JR, Westers RW.
XX	
DI	WPI: 1993-182244/22.
XX	
PT	Serine proteases derived-polypeptide(s) and anti-peptide
PT	antibodies - for inhibiting coagulation and assaying for the
PT	presence of serine protease in fluid samples

```

RESULT 6
AAFA0178      AAP40178 standard; Protein: 456 AA.
XX
AC      AAP40178.
XX
DT      11-FEB-1992 (first entry)
DE      Part of the sequence of human factor IX.
XX
KM      Christmas disease; therapy; haemophilia; factor IX; blood clotting;
XX      diagnosis.
XX
OS      Homo sapiens.
XX
PI      Key
XX      Peptide
XX      Protein
XX      Location/Qualifiers
XX      42..456
XX
PN      MOB400560-A.
XX
PD      16-FEB-1984.
XX
PF      03-AUG-1983; 83MO-G800191.
XX
PK      06-MAY-1983; 83GB-0012491.
XX
PR      04-AUG-1982; 82GB-0022485.
XX
PA      (NATR ) NATIONAL RES DEV CORP.
XX      (BROW/) BROWNLEE G.C.
XX
PI      Brownlee G., Choo KH;
XX      NPI. 1984-045931/08.
XX      N-PSDB; AAMW01Z.
XX
PT      Recombinant DNA cloning vehicles - useful in produ. of factor IX
XX      polypeptide and of diagnostic probes for Christmas disease
XX
PS      Example: Fig 9; 71pp; English.
XX
CC      The inventors claim DNA molecules comprising part or all of the
XX      human factor IX DNA. The invention also includes cDNA derived from
XX      human factor IX RNA. Specifically claimed are: recombinant 1174.5 the
XX      plasmid present in E.coli NCIB No. 11745; the modified pMT 153
XX      recombinant DNA in which the cloning vehicle is the modified pMT 153
XX      plasmid present in E.coli NCIB No. 11747; Recombinant DNA in which
XX      the bovine factor IX DNA sequence is contained in the recombinant
XX      DNA transformed into E.coli to form a clone deposited as NCIB No.
XX      11748.
XX
SQ      Sequence 456 AA:
XX
Query Match          100.0%; Score 46; DB 5; Length 456;
Best Local Similarity 100.0%; Pval No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY      1 LVPRATCAR 9
        |||
        |||
        |||
        |||
        |||
        |||
        |||
        |||
        |||
        |||
DB      371 LVRATCAR 379
XX
RESULT 7
AAP40222
ID      AAP40222 standard; Protein: 456 AA.
XX
AC      AAP40222:
XX
DT      13-FEB-1992 (first entry)
XX
Sequence encoded by part of the sequence of human factor IX cDNA.
```

XX Haemophilia: Christmas disease; diagnosis; treatment.
 XX Homo sapiens.
 XX Key
 XX Peptide 1..41
 XX Protein /label= signal
 XX 42..456
 XX CB125409-A.
 XX
 XX 07-MAR-1984.
 XX
 XX 03-AUG-1983: 83GB-0020975.
 XX
 XX 16-MAY-1983: 83GB-0012490.
 XX
 XX 04-AUG-1982: 82GB-0022486.
 XX
 XX 03-AUG-1983: 83GB-0020975.
 XX
 XX (NATR) NATIONAL RES DEV CORP.
 XX
 XX Brownlee GG, Choo KH;
 XX
 XX WPI: 1984-057898/10.
 XX
 XX N-PSDB: AAN40177.
 XX
 XX Prodn. of artificial human factor IX - by use of recombinant DNA
 XX sequences for host transformation and cultivation
 XX
 XX Example: Fig 9: 49pp: English.
 XX
 XX The inventors claim a recombinant DNA having a human factor IX
 XX sequence pref. at least 50 nucleotides long, esp. 75-27000
 XX nucleotides. A cloning vector, hereinafter referred to as also claimed,
 XX comprising a cloning site, is also claimed.
 XX The foregoing sequence of nucleotides is the sequence of a modified
 XX The human factor IX genome. The cloning vehicle may be a modified
 XX par 153 plasmid. Also claimed is a labeled diagnostic probe
 XX comprising a DNA molecule having a single- or double-stranded probe
 XX sequence of 15 to 10000 nucleotides long Factor IX DNA sequence.
 XX
 XX Sequence 456 AA:
 XX
 XX Query Match 100.0%; Score 46; DB 5; Length 456;
 XX Best Local Similarity 100.0%; Pred. No. 1.1;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 0Y 1 LVDRATCLR 9
 XX | | | | | | | | | |
 XX DB 371 LVDRATCLR 379

XX (TRAN-) TRANSGENE SA.
 XX (JAVE/) JAVE M.
 XX
 XX Jaye M, De la Salle H, Tolstoshev P, Lecoq JP;
 XX
 XX WPI: 1985-302926/48.
 XX
 XX N-PSDB: AAN50362.
 XX
 XX Cloning and expression vector for factor nine - for transforming
 XX bacteria, yeast or mammalian cells
 XX
 XX Disclosure: Fig 3, 58pp: French.
 XX
 XX The inventors claim a vector which comprises a DNA SO coding for
 XX Factor IX (FIX) and elements providing expression of this SO in host
 XX cells. The prepn. of DNA FIX is as follows. Polyadenylated mRNA is
 XX isolated from human liver, used to make ss-DNA which is digested
 XX with SI and the fragments over 1 kb cloned in pBR 322. Clones are
 XX selected by hybridisation with a 32-mer probe corresp. to the bovine
 XX factor IX SQ (AAN50362).
 XX
 XX Sequence 461 AA:
 XX
 XX Query Match 100.0%; Score 46; DB 6; Length 461;
 XX Best Local Similarity 100.0%; Pred. No. 1.1;
 XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 0Y 1 LVDRATCLR 9
 XX | | | | | | | | | |
 XX DB 376 LVDRATCLR 384

RESULT 9
 AAP50302
 ID AAP50302 standard; Protein: 461 AA.
 XX
 XX AAP50302:
 XX
 XX 03-SEP-1991 (first entry)
 XX
 XX Sequence of human factor IX.
 XX
 XX Blood clotting; haemophilia B; vaccinia vector; compox vector.
 XX
 XX Homo sapiens.
 XX
 XX W08503376-A.
 XX
 XX 05-DEC-1985.
 XX
 XX 29-MAY-1985: 85NO-1104408.
 XX
 XX 05-OCT-1984: 84FR-0015294.
 XX
 XX 22-MAY-1984: 84FR-0007959.
 XX
 XX (TRAN-) TRANSGENE SA.
 XX (DSAL/) DE LA SALLE H.
 XX
 XX De la Salle H, Drillion R, Altunbuerger W, Tolstoshev P, Lecoq JP.
 XX
 XX N-PSDB: AAN50351.
 XX
 XX Vector for expressing factor 9 in vertebrate cells - conty.
 XX poxvirus genome in which factor 9 gene is inserted
 XX
 XX Example: Fig 2: 59pp: French.
 XX
 XX The factor IX coding SQ (AAN50351) is first inserted into a non-
 XX essential region of viral DNA, cloned in a bacterial plasmid. A
 XX double-reciprocal recombination procedure is then used to transfer
 XX the insert into the viral genome where it is propagated and

```

CC expressed.
XX
SQ Sequence 461 AA:
Query Match 100.0%; Score 46; DB 6; Length 461:
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 10
AAP50019
ID AAP50019 standard; Protein: 461 AA.
XX
XX AAP50019:
XX
XX 06-SEP-1991 (first entry)
XX
XX DE Sequence of human factor IX.
XX
XX Haemophilia B; blood clotting; factor IX.
XX
XX Homo sapiens.
XX
XX EPI62782-A.
XX
XX 27-OCT-1985.
XX
XX 21-MAY-1985; 85EP-0400999.
XX
XX 05-OCT-1984; 84FR-0015394.
XX
XX 22-MAY-1984; 84FR-0007559.
XX
XX (TRAN-) TRANSGENE SA.
XX
XX de la Salle R, Drillion R, Altonburger W, Tolloshev P;
XX Lecocq JP;
XX WPI: 1985-298122/48.
XX N-PSDB; AAN50049.
XX
XX Vector for expressing factor 9 in vertebrate cells - comprises
XX pox:virus genome contg. factor 9 gene
XX
XX Example: Flg 2; 56pp; French.
XX
XX The inventors claim vectors for expressing factor IX, or an
XX analogous protein in vertebrate cells which consist of the genome of
XX poxvirus into which a gene coding for factor IX or an analogous
XX protein has been inserted. Cultivation of infected cells produces
XX very pure factor IX (a deficiency of which causes haemophilia B).
XX
XX Sequence 461 AA:
Query Match 100.0%; Score 46; DB 6; Length 461:
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 11
AAR05393
ID AAR05393 standard; Protein: 461 AA.
XX
XX AAR05393:
XX
XX 12-OCT-1990 (first entry)

```

```

XX
XX Mutant human factor IX (FIX).
DE
XX Factor IX; haemophilia B; vaccine; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifier
XX mutation replace(44,Pro)
XX mutation replace(47,Tyr)
XX
XX EPI37012-A.
XX
XX 13-JUN-1990.
XX
XX 07-NOV-1989; 89EP-0403065.
XX
XX 09-NOV-1988; 88FR-0014635.
XX
XX (TRAN-) TRANSGENE SA.
XX
XX Meullien P;
XX
XX WPI: 1990-180758/24.
XX
XX N-PSDB; AAO05393.
XX
XX DNA coding for human factor IX -
XX with mutation in pro coding sequence.
XX
XX Disclosure; ; P; French.
XX
XX The slightly altered gene product has a higher ratio of FIX
XX activity to FIX yield.
XX
XX Sequence 461 AA:
SQ
Query Match 100.0%; Score 46; DB 11; Length 461:
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

RESULT 12
AAM40284
ID AAM40284 standard; Protein: 461 AA.
XX
XX AAM40284:
XX
XX 16-JUN-1998 (first entry)
XX
XX DE Human Factor IX protease.
XX
XX Factor X; factor IX; serine protease activity; catalytic domain; ZAD;
XX zymogen activating domain; epidermal growth factor-like domain; EGF;
XX signal; signal; signal; signal; signal; signal; signal; signal;
XX detection; drug modelling; restriction protease.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..46
XX 17061 /label= signal
XX 17061 /label= Factor IX
XX 130..226 /label= EGF2
XX 227..324 /label= catalytic domain
XX
XX W09747737-A1.

```

XX 18-DEC-1997.
 XX 11-JUN-1997. 97MO-EP03027.
 XX 06-JUL-1996. 96EP-0110959.
 XX 11-JUN-1996. 96EP-0109288.
 XX 22-JUN-1996. 96EP-0110109.
 XX (BOE) BOEHRINGER MANNHEIM GMBH.
 XX Hopfner K. Kopeckzi E.
 XX MPI: 1998-052304/05.
 XX N-PSDB: AAY10463.
 XX
 XX Non-glycosylated, truncated forms of factor IX family protein with
 XX serine protease activity - used to screen for specific modulators
 XX and to assay factor IXa
 XX
 XX Disclosure: Fig 4; 49pp; German.
 XX
 XX This sequence represents a human factor IX protease. This protein is used
 XX in the construction of a novel non-glycosylated protein and truncated
 XX and zymogen forms of this protein, which have serine protease activity.
 XX The protein is composed of various domains from a factor IX family
 XX protein, namely a catalytic domain (CD), N-terminally bound ACG and/or
 XX B2 domain (B2) and a B1 domain (B1). Such proteins
 XX are used to identify activators/inhibitors of factor IX family proteins
 XX (potentially useful as regulators of coagulation, fibrinolysis and
 XX homeostasis). The protein in zymogen form is also useful in assays for
 XX detecting factor IXa activity in aqueous solution (specifically in body
 XX fluids). The protein can be used to produce catalytically active protease
 XX variants of inhibitor factor IXa specific protease variants and
 XX CC as restriction processes in biotechnology. These truncated proteins
 XX CC have the same specificity as factor IX family proteases and can be
 XX CC produced in prokaryotes in a form that allows production of active enzyme
 XX CC by conversion to native form and enzymatic cleavage.
 XX
 XX Sequence 461 AA:
 SO
 Query Match 100.0%; Score 46; DB 19; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LVDRATCLR 9
 DB 376 LVDRATCLR 384
 RESULT 13
 AAY97295
 ID AAY97295 standard; Protein: 461 AA.
 XX
 XX AAY97295:
 03-JAN-2001 (first entry)
 DE Human clotting factor IX.
 XX
 XX Recombinant DNA, gene therapy; hormone responsive element;
 KW transgene; HRE; haemophilia; clotting factor IX; vaccine;
 KW regulation; breast cancer; ovarian cancer; prostate cancer;
 KW von Willebrand disease; cystic fibrosis; hormone; receptor; human;
 XX blood.
 XX
 XX Homo sapiens.
 OS
 XX MO200049147-A1.
 PN
 XX 24-NOV-2000.
 PD

PF 18-FEB-2000; 2000MO-EP01368.
 XX 19-FEB-1999. 99DE-1007099.
 XX 19-FEB-1999. 9905-0120840.
 XX (THER-) THERAGENE BIOMEDICAL LAB GMBH.
 XX
 XX Hauser-Funke C.
 XX MPI: 2000-549273/50.
 XX N-PSDB: AAY3835.
 XX
 XX Novel nucleic acid construct useful in gene therapy comprising an
 XX hormone responsive element and transgene in which the hormone
 XX responsive element is not functionally linked to the transgene
 XX
 XX Disclosure: Page 81-83; 100pp; English.
 XX
 XX New nucleic acid constructs are described which comprise an hormone
 XX responsive element (HRE) and a transgene (T). Alternatively the
 XX HRE and T are not functionally linked to each other, where one of the HREs
 XX is not functionally linked to the transgene.
 XX The constructs can be used to up-regulate or down-regulate target
 XX genes and for the delivery of vaccines. The constructs preferably
 XX comprise a transgene which encodes a protein which is lacking in a
 XX variety of genetic disorders or involved in conditions related in
 XX inappropriate responses to hormones, for example hormone-dependent
 XX cancer. The transgene may also be used to replace a defective gene
 XX resulting in such genetic disorders as haemophilia, von Willebrand
 XX disease, and cystic fibrosis. Vectors comprising these constructs
 XX where the transgene is human clotting factor IX can be used for
 XX treating blood clotting disorders such as haemophilia A or B on
 XX administration to an organism or to a cellular system. The constructs
 XX comprise a hormone responsive element (HRE) and a transgene (T). The
 XX transgene encodes a clotting factor such as clotting factor IX. The
 XX advantage of this system is that the hormone-hormone receptor complex
 XX contains a hormone receptor that becomes activated after binding of
 XX its specific hormone. The hormone receptor in the activated state is
 XX able to recognise and bind to its specific hormone responsive
 XX element. The presence of the hormone responsive element on the
 XX nucleic acid carrying a transgene encourages binding of a hormone
 XX receptor to the hormone responsive element. The hormone receptor
 XX acts as a link between the nucleic acid carrying the transgene and the
 XX hormone known to interact with the cell membrane.
 XX
 XX Sequence 461 AA:
 SO
 Query Match 100.0%; Score 46; DB 21; Length 461;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 LVDRATCLR 9
 DB 376 LVDRATCLR 384
 RESULT 14
 AAE10828
 ID AAE10828 standard; Protein: 461 AA.
 XX
 XX AAE10828:
 18-DEC-2001 (first entry)
 DE Human wild-type factor IX protein.
 XX
 XX Human wild-type factor IX protein.
 KW human; haemostatic; coagulant; blood clotting factor; factor VIII;
 KW factor IX; therapy; haemophilia A.
 XX
 XX Homo sapiens.
 OS
 XX MO200170968-A2.
 PN

XX 27-SEP-2001.
 PD 21-MAR-2001: 2001MO-EP03220.
 PR 22-MAR-2000: 2000EP-0106225.
 PR 08-MAY-2000: 2000US-203249P.
 PA (CCTA-1) OCTAGENE CMBH.
 XX Hauser C, Hoerster A, Schroeder C, Lehnerer M:
 DR WPL: 2001-590175/66.
 DR N-PSDB: AAD18173.
 PT Recombinantly producing human blood coagulation factors VIII and IX for
 PT use in treating hemophilia -
 XX Claim 21: Page 73-74; 104P: English.
 XX The present invention relates to an improved method for the production
 CC of recombinant human blood clotting factors especially for the vitrifying
 CC IX, using an immortalised cell line stably expressing viral transcription
 CC activators and carrying a vector comprising a promoter and a sequence
 CC encoding the blood coagulation factor. The factor VIII mtein or a gene
 CC transfer vector is used in the preparation of agents for treating
 CC hemophilia, especially haemophilia A. The present sequence is human
 XX wild-type factor IX protein encoded by vector pTFcF3bHyg.
 XX Sequence 461 AA:
 S0
 YY 1 UNPACTLR 9
 YY | | | | | | | | | |
 DB 376 LDRATCLR 384
 RESULT 15
 AAB60281
 XD AAB60281 standard: Protein: 461 AA.
 AC AAB60281:
 XX 30-MAR-2001 (first entry)
 DT Human factor IX (hFIX) protein.
 DE
 XX Age-related gene regulation: liver-specific; gene expression;
 KM age-related; hFIX; hF3; hF3; age-regulatable expression construct;
 KM antithrombin; hFIX; hF3; hF3; hF3; hF3; hF3; hF3; hF3; hF3; hF3;
 KM diseases; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
 KM osteoarthritis; dementia.
 XX home sapiens.
 OS
 XX MO200075279-A2.
 PA 14-DEC-2000.
 PF 06-JUN-2000: 2000MO-US15728.
 PR 09-JUN-1999: 99US-0328925.
 PA (UNML) UNIV MICHIGAN.
 PA Kirschl K, Kirschl S:
 PR WPI: 2001-061708/07.
 DR N-PSDB: AAF54018.
 XX

PT New regulatory elements that control age-related gene expression
PT useful in gene therapy and for reducing factor IX expression -
XX
PS Disclosure: Fig 8A-E; 225pp; English.
XX

The invention relates to nucleic acid sequences which regulate gene expression in an age-related manner and/or in a liver-specific manner. The invention identifies regions C (nuc) of the human factor IX (*FIX*) gene, and a region of the human protein C (nuc) gene, which are age-related regulatory sites. These age-related regulatory sequences are 5' UTR (at position 2164-2165 of APMF408) and 3' UTR (at position 34383-3555 of APMF5018), respectively. These elements act synergistically to increase *FIX* levels over the lifespan of an individual; however, they can independently exert effects on *FIX* mRNA in an age-related manner, with A57, acting to stabilise *FIX* mRNA, and A53, acting to increase *FIX* mRNA levels, over time. A57 also directs liver-specific expression. The hpc gene age-related regulatory sequence is found in the 5' UTR (AAAF50181), and contains two PEA-3 (polyoma virus activator 3) elements of the GAGAAGAA... and 3'-GACAGAG-3'. The age-related sequences of the GAGAAGAA... and 3'-GACAGAG-3' may be used in the construction of recombinant expression vectors for the expression of a desired sequence in an age-related fashion in a host cell. Preferred target genes for expression in such age-regulatable expression vectors include those encoding proteins involved in blood coagulation (e.g., the pro-coagulants factor XI and factor VIII, and the anti-coagulants protein C and antithrombin III), human alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as luciferase. Preferred promoters for use in such age-regulatable constructs include the human factor IX promoter. The invention provides the T3 promoter, which is useful for directing the expression of an invention may be used in gene therapy to provide age-related and/or liver-specific expression of target genes. Age-regulatable constructs may be used in the treatment of such age-related conditions such as thrombosis, cardiovascular disease, diabetes, Alzheimer's disease, Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia. Specifically, they may be used to express factor IX antisense RNA in the treatment of thrombotic conditions associated with the natural plasma levels of factor IX. The present invention provides methods for identifying animals that contain vectors of the invention are useful as models of these diseases, in screening for potential therapeutic agents and for studying normal processes such as ageing and gene expression. Fragments and homologues of age-related regulatory sequences, are useful as probes to detect, isolate or identify other such sequences in samples. The present sequence represents *FIX*.

```

SQ      Sequence 461 AA;
Query Match      100.0%; Score 46; DB 23; Length 461;
Local Similarity 100.0%; P-Val. No. 1;
Matches 9; Conserved 0; Mismatches 0; Indels 0; Gaps 0
OY      1 EVDRATCLR 9
Db      376 EVDRATCLR 384

```

```
Search completed: May 6, 2003, 15:01:01
Job time : 42.95 secs
```


GenCore version 5.1.4-PS4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:30 Search time 14.4 Seconds
(without alignments)
18,389 Million cell updates/sec

Title: US-09-851-422b-1

Perfect score: 46

Sequence: 1 LYRATCLR 9

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Minimum Match 100%

Listing first 45 summaries

Database: Issued Patents, AA:*

1: /cgn2_6/pdata1/1aa/5A.COMB.pep:*

2: /cgn2_6/pdata1/1aa/5B.COMB.pep:*

3: /cgn2_6/pdata1/1aa/6A.COMB.pep:*

4: /cgn2_6/pdata1/1aa/6B.COMB.pep:*

5: /cgn2_6/pdata1/1aa/6C.COMB.pep:*

6: /cgn2_6/pdata1/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	40	US-08-797-842-11	Sequence 11, Appl
2	46	100.0	40	US-08-797-842-12	Sequence 12, Appl
3	46	100.0	235	US-08-944-483-48	Sequence 48, Appl
4	46	100.0	415	US-08-285-411-2	Sequence 1, Appl
5	46	100.0	415	US-08-985-471-2	Sequence 1, Appl
6	46	100.0	415	US-08-766-288-1	Sequence 1, Appl
7	46	100.0	415	PCT-US92-10242-2	Sequence 2, Appl
8	46	100.0	461	US-08-742-877-2	Sequence 2, Appl
9	46	100.0	461	US-09-053-871A-21	Sequence 2, Appl
10	46	100.0	461	US-08-797-842-11	Sequence 11, Appl
11	46	100.0	461	US-08-797-842-12	Sequence 12, Appl
12	46	100.0	461	US-08-797-842-13	Sequence 13, Appl
13	46	100.0	461	US-08-797-842-14	Sequence 14, Appl
14	46	100.0	461	US-08-797-842-15	Sequence 15, Appl
15	46	100.0	461	US-08-797-842-16	Sequence 16, Appl
16	46	100.0	461	US-08-797-842-17	Sequence 17, Appl
17	46	100.0	461	US-08-797-842-18	Sequence 18, Appl
18	46	100.0	461	US-08-797-842-19	Sequence 19, Appl
19	46	100.0	461	US-08-797-842-20	Sequence 20, Appl
20	46	100.0	461	US-08-797-842-21	Sequence 21, Appl
21	46	100.0	461	US-08-797-842-22	Sequence 22, Appl
22	46	100.0	461	US-08-797-842-23	Sequence 23, Appl
23	46	100.0	461	US-08-797-842-24	Sequence 24, Appl
24	46	100.0	461	US-08-797-842-25	Sequence 25, Appl
25	46	100.0	461	US-08-797-842-26	Sequence 26, Appl
26	46	100.0	461	US-08-797-842-27	Sequence 27, Appl
27	46	100.0	461	US-08-797-842-28	Sequence 28, Appl
28	46	100.0	461	US-08-797-842-29	Sequence 29, Appl
29	46	100.0	461	US-08-797-842-30	Sequence 30, Appl
30	46	100.0	461	US-08-797-842-31	Sequence 31, Appl
31	46	100.0	461	US-08-797-842-32	Sequence 32, Appl
32	46	100.0	461	US-08-797-842-33	Sequence 33, Appl
33	46	100.0	461	US-08-797-842-34	Sequence 34, Appl
34	46	100.0	461	US-08-797-842-35	Sequence 35, Appl
35	46	100.0	461	US-08-797-842-36	Sequence 36, Appl
36	46	100.0	461	US-08-797-842-37	Sequence 37, Appl
37	46	100.0	461	US-08-797-842-38	Sequence 38, Appl
38	46	100.0	461	US-08-797-842-39	Sequence 39, Appl
39	46	100.0	461	US-08-797-842-40	Sequence 40, Appl
40	46	100.0	461	US-08-797-842-41	Sequence 41, Appl
41	46	100.0	461	US-08-797-842-42	Sequence 42, Appl
42	46	100.0	461	US-08-797-842-43	Sequence 43, Appl
43	46	100.0	461	US-08-797-842-44	Sequence 44, Appl
44	46	100.0	461	US-08-797-842-45	Sequence 45, Appl

28	32	69.6	547	1	US-08-314-362-1	Sequence 1, Appl
29	32	69.6	547	1	US-08-473-981A-6	Sequence 6, Appl
30	32	69.6	547	1	US-08-433-010-1	Sequence 1, Appl
31	32	69.6	547	1	US-08-482-882-1	Sequence 1, Appl
32	32	69.6	547	1	US-08-487-1130-1	Sequence 1, Appl
33	32	69.6	547	2	US-08-487-1130-1	Sequence 1, Appl
34	32	69.6	547	2	US-08-473-503-1	Sequence 1, Appl
35	32	69.6	547	2	US-08-483-932-1	Sequence 1, Appl
36	32	69.6	547	2	US-08-474-087-6	Sequence 6, Appl
37	32	69.6	547	2	US-08-720-420A-1	Sequence 1, Appl
38	32	69.6	547	3	US-08-714-017-1	Sequence 1, Appl
39	32	69.6	547	3	US-08-853-780-1	Sequence 1, Appl
40	32	69.6	547	3	US-08-853-780-1	Sequence 1, Appl
41	32	69.6	547	4	US-08-286-749-1	Sequence 1, Appl
42	31	67.4	77	4	US-09-134-001C-3500	Sequence 3500, Ap
43	31	67.4	222	6	5223425-6	Patent No. 5223425
44	31	67.4	223	4	US-09-516-914-13	Sequence 13, Appl
45	31	67.4	846	1	US-08-149-103-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-797-842-11
Patent No. 5932706
Applicant: Mettens, Koenraad et al
Title of Invention: Antibodies specific for a haemostatic protein,
Title of Invention: their use for isolating intact protein, haemostatic compo
Title of Invention: of proteolytic cleavage products of the protein
Number of Sequences: 16
Correspondence Address:
Address: Michaelson and Wallace
Street: 109 Little Center, 328 Newman Springs
City: Red Bank
State: New Jersey
ZIP: 07701
Country: USA
Computer Readable Form:
Medium Type: 3 1/2" 1.44 MByte IBM compatible diskette
Operating System: MS-DOS, Windows for Workgroups 3.11
Software: Microsoft Word for Windows 6
Current Application Data:
Application Number: US/08/797,842
Filing Date: 10-Feb-1997
Classification: 530
Priority Date: 10-Feb-1997
Application Number: 08/381,891
Filing Date: February 8, 1995
Attorney/Agent Information:
Name: Michaelson, Peter L.
Registration Number: 30090
Reference/Docket Number: Sticking-5
Telecommunication Information:
Telephone: (908)530-6871
Telefax: (908)530-6884
Information for SEQ ID NO: 11:
Sequence Characteristics:
Length: 40 amino acids
Type: amino acid
Strandedness: unknown
Molecular Weight: unknown
Molecular Weight: unknown
Hypothetical: NO
US-08-797-842-11
Query Match 100.0%, Score 46; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDATCLR 9
DB 23 LVDATCLR 31

RESULT 2

US-08-797-842-12
Sequence 12, Application US/08/97842
Patent No. 593706
GENERAL INFORMATION:
TITLE OF INVENTION: Monoclonal antibodies specific for a haemostatic protein.
TITLE OF INVENTION: Their use for isolating intact protein, haemostatic composite
TITLE OF INVENTION: of proteolytic cleavage products of the protein
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michaelson and Wallace
STREET: 1933 West 19th Street, Suite 200, Center, 328 Newman Springs
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2" 1.44 Mbyte IBM compatible diskette
OPERATING SYSTEM: MS-DOS Windows for Workgroups 3.11
SOFTWARE: Microsoft Word for Windows 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,842
FILING DATE: 10-Feb-1997
CLASSIFICATION: C30
PRIORITY CLAIM: YES
APPLICATION NUMBER: A:08/781,891
FILING DATE: February 8, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Stliching-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)530-6571
TELEFAX: (908)530-6581
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
MOLECULE TYPE: unknown
HYPOHETICAL: NO

US-08-797-842-12
Query Match 100.0%; Score 46; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDATCLR 9
DB 12 LVDATCLR 20

RESULT 3
US-08-944-483-48
Sequence 48, Application US/08/944483
Patent No. 6232456
GENERAL INFORMATION:
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIDMAN, PAULA N.
APPLICANT: GRAMADOS, EDWARD N.
APPLICANT: KILASS, MICHAEL R.
APPLICANT: ROSSBELL, JOHN C.

APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTISO for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,483
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6103.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623

INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 235 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 62324566

US-08-944-483-48

Query Match 100.0%; Score 46; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDATCLR 9
DB 150 LVDATCLR 158

RESULT 4
US-08-073-531B-1
Sequence 1, Application US/08/073531B
Patent No. 5621039

GENERAL INFORMATION:
APPLICANT: Hallahan, et al.
APPLICANT: INVENTOR IX - Polymeric Conjugates
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: GALIANO & BURKE
STREET: 300 Babro Drive
CITY: Haupauge
STATE: New York
COUNTRY: USA
ZIP: 11783
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS 6.0
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/073,531B

```

? FILING DATE: June 8, 1993
? CLASSIFICATION: 525
? PUBLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: GALICANO & BURKE
? REGISTRATION NUMBER: 30,735
? REFERENCE/DOCKET NUMBER: 128-7
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-582-6191
? TELEFAX: 415-582-6191
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 415 Amino Acids
? TYPE: Amino Acid
? STANDARDS: Single
? MODIFICATION: Unknown to applicant
? HYPOTHETICAL:
? ANTI-SENSE:
? ORIGINAL SOURCE:
? ORGANISM:
? INDIVIDUAL ISOLATE:
? CELL TYPE:
? IMMEDIATE SOURCE:
? CLONE:
? PUBLICATION INFORMATION:
? AUTHORS:
? TITLE:
? JOURNAL:
? VOLUME:
? PAGES:
? DATE:
? RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 415.
US-08-073-531B-1
?
? Query Match 100.0%; Score 46; DB 1; Length 415;
? Best Local Similarity 100.0%; Pred. No. 0.21;
? Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
? Oy 1 LVDRATCLR 9
? Db 330 LVDRATCLR 338

```

```

? APPLICATION NUMBER: US/08/295,411
? FILING DATE: 22-AUG-1994
? CLASSIFICATION: 530
? PUBLICATION DATA:
? PRIOR APPLICATION NUMBER: US 07/793,969
? FILING DATE: 18 NOV-1991
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: Filling, Thomas
? REGISTRATION NUMBER: 34,163
? REFERENCE/DOCKET NUMBER: TSH1263.0C1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-582-6191
? TELEFAX: 415-582-6191
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 415 amino acids
? TYPE: amino acid
? MODIFICATION:
? HYPOTHETICAL:
? ANTI-SENSE: NO
? FEATURE:
? NAME/KEY: Region
? LOCATION: 1..145
? OTHER INFORMATION: /note="Factor IX light chain"
? NAME/KEY: Region
? LOCATION: 146..180
? OTHER INFORMATION: /note="Factor IX Activation"
? OTHER INFORMATION: Peptide"
? FEATURE:
? NAME/KEY: Region
? LOCATION: 181..415
? OTHER INFORMATION: /note="Factor IX heavy chain"
US-08-295-411-2
?
? Query Match 100.0%; Score 46; DB 1; Length 415;
? Best Local Similarity 100.0%; Pred. No. 0.21;
? Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
? Oy 1 LVDRATCLR 9
? Db 330 LVDRATCLR 338

```

```

? RESULT 5
? US-08-295-411-2
? Sequence 2, Application US/08295411
? Patent No. 5679659
? GENERAL INFORMATION:
? APPLICANT: Griffin, John H.
? APPLICANT: Westers, Rolf M.
? TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
? TITLE OF INVENTION: Anti-Peptide Anticoagulants, Systems and Therapeutic Methods
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Office of Patent Counsel, The Scripps
? STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
? CITY: La Jolla
? STATE: CA
? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:

```

```

? RESULT 6
? US-08-955-471-2
? Sequence 2, Application US/08955471
? Patent No. 5968751
? GENERAL INFORMATION:
? APPLICANT: Griffin, John H.
? APPLICANT: Westers, Rolf M.
? TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
? TITLE OF INVENTION: Anti-Peptide Anticoagulants, Systems and Therapeutic Methods
? NUMBER OF SEQUENCES: 10
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Office of Patent Counsel, The Scripps
? STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8
? CITY: La Jolla
? STATE: CA
? COUNTRY: USA
? ZIP: 92037
? COMPUTER READABLE FORM:
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? FILING DATE:

```

```

CLASSIFICATION:
PRIOR APPLICATION DATA: 08/295,411
FILING DATE: 08/295,411
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSH1263, OC1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-6312
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYDROPHILIC: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note="Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note="Factor IX Activation"
FEATURE:
NAME/KEY: Peptide
OTHER INFORMATION:
NAME/KEY: Region
LOCATION: 181..415
OTHER INFORMATION: /note="Factor IX Heavy Chain"
US-08-955-471.2
Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LVDAATCLR 9
Db 330 LVDAATCLR 338

```

```

REGISTRATION NUMBER: 30,735
REFERENCE/DOCKET NUMBER: 128-7 (DIV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (216) 562-6161
TELEFAX: (216) 562-6161
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 Amino Acids
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown to applicant
MOLECULE TYPE:
HYDROPHILIC:
ANTI-SENSE:
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE:
CELL TYPE:
IMMEDIATE SOURCE:
CROSS:
CLONING:
PUBLICATION INFORMATION:
AUTHORS:
TITLE: JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 415.
US-08-766-288-1
Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LVDAATCLR 9
Db 330 LVDAATCLR 338

```

```

RESULT 7
US-08-766-288-1
Sequence 1, Application US/08766288
GENERAL INFORMATION:
APPLICANT: Hallahan, et al.
TITLE OF INVENTION: Factor IX - Polymeric Conjugates
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: GALICANO & BURKE
CITY: HANNOU RABO Drive
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: NCH/Doc 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,288
FILING DATE:
CLASSIFICATION: 525
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,531
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: GALICANO & BURKE

```

```

RESULT 8
FCT-US92-10242-2
Sequence 2, Application PC/TUS9210242
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESS: Research Institute
STREET: 10666 North Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PC/US92/10242
FILING DATE: 19921118
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,989
FILING DATE: 18-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas

```

```

REGISTRATION NUMBER: 34,163
PATENT NUMBER: US/08742877
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2837
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: AMINO ACID
SOURCE: HUMAN
MOLECULE TYPE: linear
HYDROPHILIC: NO
HYDROPHOBIC: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note="Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note="Factor IX Activation"
FEATURE:
NAME/KEY: Peptide
LOCATION: 181..415
OTHER INFORMATION: /note="Factor IX Heavy Chain"
PCT-US92/10442-2
Query Match 100.0%; Score 46; DB 5; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDATCFLR 9
DB 330 LVDATCFLR 338

```

```

REGISTRATION NUMBER: 34,331
PATENT NUMBER: US/0470001/REF
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
SOURCE: HUMAN
MOLECULE TYPE: protein
US-08-742-877-2
Query Match 100.0%; Score 46; DB 3; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDATCFLR 9
DB 376 LVDATCFLR 384

```

```

RESULT 9
US-08-742-877-2
Sequence 2, Application US/08742877
Patent No. 606380
GENERAL INFORMATION:
APPLICANT: CLARK, Anthony J.
TITLE OF INVENTION: DNA SEQUENCES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
STATE: WASHINGTON
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,877
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9408717.8
FILING DATE: 02-NOV-1994
ATTORNEY/DOEN INFORMATION:
NAME: FLEISHER, RAY E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0633, 0470001/REF
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 461 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
SOURCE: HUMAN
MOLECULE TYPE: protein
US-08-742-877-2
Query Match 100.0%; Score 46; DB 3; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDATCFLR 9
DB 376 LVDATCFLR 384

```

```

RESULT 10
US-09-053-871A-21
Sequence 21, Application US/09053871A
Patent No. 631995
GENERAL INFORMATION:
APPLICANT: Plinsky, David J.
APPLICANT: Stern, David
APPLICANT: Rose, Eric
APPLICANT: Solomon, Robert A.
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: METHODS FOR TREATING AN ISCHEMIC DISORDER AND IMPROVING
THE OUTCOME OF STROKE
CURRENT FILING DATE: 1998-04-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
LENGTH: 461
SOURCE: HUMAN
ORGANISM: Homo Sapien
US-09-053-871A-21
Query Match 100.0%; Score 46; DB 4; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDATCFLR 9
DB 376 LVDATCFLR 384

```

```

RESULT 11
US21070-2
Patent No. 5521070
GENERAL INFORMATION:
TITLE OF INVENTION: DNA SEQUENCE CODING FOR HUMAN FACTOR
IX OR A SIMILAR PROTEIN, EXPRESSION VECTOR, TRANSFORMED CELLS,
METHOD FOR PREPARING FACTOR IX AND CORRESPONDING PRODUCTS OBTAINED
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/209,489
FILING DATE: 03-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 970,966
FILING DATE: 03-NOV-1992
APPLICATION NUMBER: 433,276
FILING DATE: 08-NOV-1989
SEQ ID NO:2:
LENGTH: 461
US21070-2
Query Match 100.0%; Score 46; DB 5; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDATCFLR 9
DB 330 LVDATCFLR 338

```

DB 376 LVDRATCLR 384

RESULT 12

US-08-944-483-44

Sequence 44, Application US/08944483

Patent No. 623456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLETTIS, TRACEY L.

APPLICANT: CRIVELLO, JAMES M.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLAS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: TROPE, STEVEN D.

TITLE OF INVENTION: NOVEL STRAIN PROTEASE REAGENTS

TITLE OF INVENTION: NOVEL STRAIN PROTEASE REAGENTS FOR DETECTING AND TREATING DISEASES

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSER: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

COUNTRY: IL USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

ATTILING DATE:

NAME/REG. INFO:

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183 US 01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/935-2623

TELEX:

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 228 amino acids

TYPE: amino acids

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: NO. 62344566

US-08-944-483-44

Query Match

Best Local Similarity 66.7%; DB 4; Length 228;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 148 LVDRATCLR 156

US-08-944-483-44

RESULT 13

5223425-5

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/277,963

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 5

LENGTH: 238

Query Match

Best Local Similarity 66.7%; DB 6; Length 238;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 158 LVDRATCLR 166

US-08-944-483-44

RESULT 14

5223425-4

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/277,963

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 5

LENGTH: 250

Query Match

Best Local Similarity 66.7%; DB 6; Length 250;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 170 LVDRATCLR 178

US-08-944-483-44

RESULT 15

5223425-8

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/277,963

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 8

LENGTH: 253

Query Match

Best Local Similarity 66.7%; DB 6; Length 253;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 173 LVDRATCLR 181

Mon May 12 15:29:54 2003

us-09-851-422b-1.ra1

Page 7

Search completed: May 6, 2003, 15:04:05
Job time : 16.4 secs

THIS PAGE BLANK (00070)

PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30/60/734,687
PRIOR FILING DATE: 2000-09-21/60/734,687
PRIOR FILING DATE: 2000-09-21/60/734,687
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
NOT NAME: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO: 18100
LENGTH: 18100
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL033403.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
OTHER INFORMATION: EXPRESSED IN HUMAN LIVER, SIGNAL = 0.0076
OTHER INFORMATION: SWISSPROT HIT: P00740, EVALUE 1.00e-107
US-09-864-761-46600

Query Match 100.0% Score 46; DB 10; Length 181;
Best Local Similarity 100.0%; Pred No. 0.18; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRACTLR 9
DB 96 LVDRACTLR 104

RESULT 2
US-10-132-829-5
Sequence 351, Application US/0132829
Publication No. US20030044982A1

GENERAL INFORMATION:
APPLICANT: Hosomi, Masahiko
APPLICANT: Chien, Kenneth R
TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VIII
TITLE OF INVENTION: vltch vesicle vector
CURRENT APPLICATION NUMBER: US/10/132,829
CURRENT FILING DATE: 2002-04-25
PRIOR APPLICATION NUMBER: 60/286,314
PRIOR FILING DATE: 2001-04-25
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 5
SEQ ID NO: 561
TYPE: PRT
ORGANISM: Homo sapiens
US-10-132-829-5

Query Match 100.0% Score 46; DB 9; Length 401;
Best Local Similarity 100.0%; Pred No. 0.46; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRACTLR 9
DB 376 LVDRACTLR 384

RESULT 3
US-09-884-901-3

Sequence 3, Application US/09884901
Patent No. US20020076798A1
GENERAL INFORMATION:
APPLICANT: Miso, Carol
APPLICANT: Kay, Mark
TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods of Use
TITLE OF INVENTION: U09F-1,17396
CURRENT APPLICATION NUMBER: US/09/884,901
CURRENT FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: US 60/212,902
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.0
SEQ ID NO: 3
LENGTH: 461
TYPE: PRT
ORGANISM: Homosapien
US-09-884-901-3

Query Match 100.0% Score 46; DB 10; Length 461;
Best Local Similarity 100.0%; Pred No. 0.46; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRACTLR 9
DB 376 LVDRACTLR 384

RESULT 4
US-09-118-748-2
Sequence 2, Application US/09118748A
Patent No. US20020031799A1

GENERAL INFORMATION:
APPLICANT: Staifford, Darrel W.
APPLICANT: Chinn, Jinhua
TITLE OF INVENTION: Antithrombotic Factor with increased clotting
FILE REFERENCE: 5470-183
CURRENT APPLICATION NUMBER: US/09/118,748A
CURRENT FILING DATE: 1998-07-17
EARLIER FILING DATE: 60/053,571
EARLIER FILING DATE: 1997-07-21
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO: 2
LENGTH: 415
TYPE: PRT
ORGANISM: Homo sapiens
US-09-118-748-2

Query Match 99.1% Score 41; DB 10; Length 415;
Best Local Similarity 100.0%; Pred No. 3.8; 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRACTLR 8
DB 330 LVDRACTLR 337

RESULT 5
US-10-012-542-351
Sequence 351, Application US/10012542
Publication No. US20030044851A1

GENERAL INFORMATION:
APPLICANT: Rubin et al.
APPLICANT: Hsiao et al.
TITLE OF INVENTION: Human Secreted Proteins
FILE REFERENCE: P2029p1
CURRENT APPLICATION NUMBER: US/10/012,542
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113

```

: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
: NUMBER OF SEQ ID NOS: 532
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 1
: LENGTH: 101
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-012-542-351

Query Match
Best Local Similarity 71.7%, Score 33; DB 9; Length 101;
Matches 6; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

Cy 3 DBACCL 8
DB 33 DBACCL 38

RESULT 6
US-09-821-255-2
: Sequence 2; Application US/09861255
: Patent No. US20020081293A1
: GENERAL INFORMATION:
: APPLICANT: Michael S.C. Fung
: APPLICANT: Bill N.C. Sun
: APPLICANT: Cecily R.Y. Sun
: FILE REFERENCE: 198-2; Inhibitors of Complement Activation
: CURRENT APPLICATION NUMBER: US/09/821,255
: PRIOR FILING DATE: 2001-03-29
: PRIOR APPLICATION NUMBER: 60/075,348
: PRIOR FILING DATE: 1998-02-20
: PRIOR APPLICATION NUMBER: 09/253,689
: PRIOR FILING DATE: 1999-02-20
: PRIOR APPLICATION NUMBER: 60/075,348
: SOFTWARE: PatsSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 228
: TYPE: PRT
: ORGANISM: human
US-09-821-255-2

Query Match
Best Local Similarity 71.7%, Score 33; DB 10; Length 228;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 1 LVDATCLR 9
DB 148 VLDATCLR 156

RESULT 7
US-10-001-843-188
: Sequence 188; Application US/10001843
: Patent No. US20020132255A1
: GENERAL INFORMATION:
: APPLICANT: Salceda, Susana
: APPLICANT: Medina, Roberto
: APPLICANT: Medina, Roberto
: APPLICANT: Caffrey, Robert
: APPLICANT: Sun, Yongming
: APPLICANT: Liu, Chenguang
: APPLICANT: Turner, Leah
: TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pr
: FILE REFERENCE: DD-076
: CURRENT APPLICATION NUMBER: US/10/001,843
: PRIOR FILING DATE: 2000-11-20
: PRIOR APPLICATION NUMBER: 60/249,992
: PRIOR FILING DATE: 2000-11-20
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 188
: LENGTH: 44
```

```

: TYPE: PRT
: ORGANISM: Homo sapien
US-10-001-843-188

Query Match
Best Local Similarity 69.6%, Score 32; DB 12; Length 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Cy 1 LVDATCLR 8
DB 7 LVDATCLR 14

RESULT 8
US-09-864-761-35244
: Sequence 35244; Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Rank, David R.
: APPLICANT: Chen, Wenhong
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aeomica X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263, 6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 60/408,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Patent Sequence Listing Engine vers. 1.1
: SEQ ID NO 35244
: LENGTH: 33
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AL03378.6
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
```

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
 OTHER INFORMATION: EXPRESSED IN BT743, SIGNAL = 6.90e-01
 OTHER INFORMATION: EST_HUMAN HIT: A014284.1, EVALUATE 5.00e-31
 US-09-854-761-35244

Query Match 69.6% Score 32: DB 10: Length 83:
 Best Local Similarity 75.0% Pctd No. 38:
 Matches 6: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 10 VDRATCLR 17

RESULT 9

US-10-117-323-38
 : Sequence 38, Application US/10117323
 : Publication No. US20030054953A1
 : GENERAL INFORMATION:
 : APPLICANT: Derick E.
 : APPLICANT: Bancourt, Susan L.
 : APPLICANT: O'Sullivan, Colleen M.
 : TITLE OF INVENTION: Implantation Serine Proteinases
 : FILE REFERENCE: 03337-005
 : CURRENT FILING DATE: 05/10/117.323
 : PRIOR APPLICATION NUMBER: US/10-04-08
 : PRIOR FILING DATE: 2001-04-06/281,724
 : PRIOR APPLICATION NUMBER: US 60/294,736
 : PRIOR FILING DATE: 2001-05-30
 : PRIOR APPLICATION NUMBER: US 60/350,962
 : PRIOR FILING DATE: 2002-01-25
 : NUMBER OF SEQ ID NOS: 41
 : SEQ ID NOS: 1-41
 : SOFTWARE: PasteSeq for Windows Version 4.0
 : LENGTH: 252
 : TYPE: PRT
 : ORGANISM: Mouse
 : US-10-117-323-38

Query Match 69.6% Score 32: DB 9: Length 252:
 Best Local Similarity 67.1% Pctd No. 1e+02:
 Matches 6: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

Db 164 VDRATCLR 9
 164 VDRATCLR 172

RESULT 10

US-09-753-436-120
 : Sequence 120, Application US/09753436
 : Patent No. US20010029293A1
 : GENERAL INFORMATION:
 : APPLICANT: Gallatin, W. Michael
 : APPLICANT: Vazquez, Rosemary
 : APPLICANT: J. M. Anderson
 : NUMBER OF SEQUENCE ADDRESSES: 120
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Marshall, O'Toole, Gerstlein, Murray & Borum
 : STREET: 6300 Sears Tower, 233 South Wacker Drive
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: United States of America
 : TELEPHONE: 312-660-8000
 : COMPUTER REPRODUCIBLE FORM:
 : MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: DOS/OS/MS-DOS
 SOFTWARE: Nucleic in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/753.436
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/382,289
 FILING DATE:
 APPLICATION NUMBER: US 08/487,113
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/286,754
 FILING DATE: 05-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/102,852
 FILING DATE: 05-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/009,266
 FILING DATE: 22-JAN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/894,061
 FILING DATE: 05-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/889,724
 FILING DATE: 26-MAY-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,689
 FILING DATE: 27-JAN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams, Joseph A., Jr.
 ADDRESS: 10000 Wilshire Blvd, Suite 600, Beverly Hills, CA 90210
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 120:
 SEQUENCE CHARACTERISTICS:
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-09-753-436-120

Query Match 69.6% Score 32: DB 10: Length 434:
 Best Local Similarity 83.3% Pctd No. 2.1e+02:
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 292 VDRATCLR 297

RESULT 11

US-10-047-542-65
 : Sequence 65, Application US/10047542
 : Patent No. US20020160367A1
 : GENERAL INFORMATION:
 : APPLICANT: LARRICK, JAMES W.
 : APPLICANT: LARRICK, JAMES W.
 : APPLICANT: KNOX, JAMES W.
 : TITLE OF INVENTION: ANTI-BACTERIAL DISEASES
 : FILE REFERENCE: 030905.0004 CIP1
 : CURRENT FILING DATE: US/10/047.542
 : PRIOR APPLICATION NUMBER: PCT/US01/13932
 : PRIOR FILING DATE: 2001-10-26
 : PRIOR APPLICATION NUMBER: US/10-04-28200.298
 : PRIOR FILING DATE: 2001-04-28
 : NUMBER OF SEQ ID NOS: 101

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 65
: LENGTH: 547
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: POLYMER TYPE: protein
: NAME/KEY: Protein
: LOCATION: 30..547
: US-10-047-542-65

Query Match
Best Local Similarity 69.6%; Score 32; DB 9; Length 547;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
DB 402 IDRATC 407

RESULT 12
US-09-753-436-1
: Sequence 11, Application US/09753436
: Patent No. US2001002293A1
: GENERAL INFORMATION:
: APPLICANT: Gallatin, W. Michael
: TITLE OF INVENTION: ICAM-Related Materials and Methods
: NUMBER OF SEQUENCES: 120
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: ADDRESS: 610 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09753,436
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION NUMBER: 09/382,289
: FILING DATE:
: APPLICATION NUMBER: US 08/487,113
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/286,754
: FILING DATE: 05-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/102,852
: FILING DATE: 05-AUG-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/009,266
: FILING DATE: 22-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/894,061
: FILING DATE: 26-MAY-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/889,724
: FILING DATE: 27-JAN-1992
: PRIOR APPLICATION DATA:
: ATTORNEY/AGENT INFORMATION:
: NAME: J. T.
: REGISTRATION NUMBER: 38,659
: REFERENCE/DOCKET NUMBER: 33282
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 1:

```

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 547 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: POLYMER TYPE: protein
: NAME/KEY: Protein
: LOCATION: 30..547
: US-09-753-436-1

Query Match
Best Local Similarity 83.3%; Score 32; DB 10; Length 547;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
DB 402 IDRATC 407

RESULT 13
US-09-910-172-29
: Sequence 29, Application US/09919172
: Patent No. US20020119463A1
: GENERAL INFORMATION:
: APPLICANT: Paris, Mary
: TITLE OF INVENTION: PROSTATE CANCER MARKERS
: FILE REFERENCE: PA-0036 US
: CURRENT FILING DATE: 2000-07-30
: PRIOR FILING DATE: 2000-07-26
: NUMBER OF SEQ ID NOS: 102
: SOFTWARE: PERL Program
: SEQ ID NO: 29
: LENGTH: 2125
: TYPE: amino acid
: ORGANISM: Homo sapiens
: NAME/KEY: misc_feature
: OTHER INFORMATION: incyle ID No. US20020119463A1 3774181CD1
: US-09-919-172-29

Query Match
Best Local Similarity 69.6%; Score 32; DB 10; Length 2125;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
DB 1946 LVDRATCLR 1954.

RESULT 14
US-10-012-542-362
: Sequence 362, Application US/10012542
: Publication No. US20030044851A1
: GENERAL INFORMATION:
: APPLICANT: Rubin et al
: TITLE OF INVENTION: Human Secreted Proteins
: FILE REFERENCE: 37289
: CURRENT FILING DATE: 2001-12-12
: PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/461,325
: PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
: PRIOR APPLICATION NUMBER: 60/089,507
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089,508
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089,509
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089,510
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/090,112

```

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
 PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 NUMBER OF SEQ ID NOS: 512
 SOFTWARE: BLASTN Ver. 2.0
 SEQ ID NO 362
 LENGTH: 24
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-012-542-362

Query Match 67.4% Score 31; DB 9; Length 24;
 Best Local Similarity 75.0% Pred No. 17;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LVDRATCL 8
 DB 12 LODRASCCL 19

RESULT 15
 US-09-870-162A-13 Application US/09870162A
 Sequence 13, Patent No. US20020042118A1
 GENERAL INFORMATION: Spline
 APPLICANT: Becton Dickinson
 TITLE OF INVENTION: Phenol Induced Proteins of Thauera aromatica
 FILE REFERENCE: BCI006 US DIV
 CURRENT APPLICATION NUMBER: US/09/870,162A
 CURRENT FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: 09/516914
 INVENTOR: Becton Dickinson
 NUMBER OF SEQ ID NOS: 44
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 13
 LENGTH: 223
 TYPE: PRT
 ORGANISM: Thauera aromatica
 US-09-870-162A-13

Query Match 67.4% Score 31; DB 10; Length 223;
 Best Local Similarity 66.7% Pred No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LVDRATCL 9
 DB 46 LITRACCL 54

Search completed: May 6, 2003, 15:04:52
 Job time : 19.333 secs


```

C:Genetics:
A:Start codon: GTG
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase
Query Match 42.2% Score 54: DB 2: Length 416:
Best Local Similarity 62.5%: pred No. 3.5*
Matches 10: Conservative 3: Mismatches 3: Indels 0: Gaps 0:
OY 9 KYLDFGKILQJLIED 24
|:|||||:||||:|
Db 62 KILDFGIKILQADAE 77

RESULT 5
T05032
Protein kinase homolog P13C5.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #excl_change 11-Jun-1999
C:Accession: T05032
R:Hevan, M.; Poll, T.; Weizengger, T.; Bancroft, I.; Neues, H.W.; Mayer, K.F.X.; Sch
A:Reference number: 215395
A:Sequence source: Sequence Database, February 1998
A:Accession: T05032
A:Residue type: DNA
A:Residues: 1-421 <ENV>
A:Cross-references: EMBL:AL021711
A:Experimental source: cultivar Columbia; BMC clone P13C5
C:Genetics:
A:Start codon: GTG
A:Internal 94/3: 119/3: 151/3: 180/2: 220/3: 247/2: 283/1: 311/3: 352/2: 389/3
A:Note: P13C5.120

Query Match 42.2% Score 54: DB 2: Length 421:
Best Local Similarity 58.8%: Pred. No. 3.5:
Matches 10: Conservative 3: Mismatches 4: Indels 0: Gaps 0:
OY 9 KYLDFGKILQJLIED 25
|:||||:|:|:|
Db 258 KYADFVSKLVTKEBK 274

RESULT 6
WZ0551
gene 51 protein - human herpesvirus 3
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #excl_change 16-Jul-1999
C:Accession: G27344
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67: 1795-1816, 1986
A:Note: The complete DNA sequence of varicella-zoster virus.
A:Accession: G27344; AF27345; M01D:8630657; PMID:3018124
A:Molecule type: DNA
A:Cross-references: 1-835 <DAN>
A:Cross-references: EMBL:X04370; NID:959989; PIDN:CAA27933.1; PID:960039
C:Genetics:
A:Gene: 51
C:Superfamily: varicella-zoster virus gene 51 protein
C:Keywords: DNA binding

Query Match 41.4% Score 53: DB 1: Length 835:
Best Local Similarity 43.5%: Pred. No. 10:
Matches 10: Conservative 5: Mismatches 8: Indels 0: Gaps 0:
OY 1 GFATCTKYLDFGKILQJLIE 33
|:|||||:||||:|
Db 225 GFSGCTCTILKQIDLVKIK 247

RESULT 7
D69053
Probable iron-sulfur protein MTH40 [similarity] - Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum

```


Mon May 12 15:29:53 2003

Job time : 18.9412 secs

us-09-851-422a-8.rpt

Page 5

THIS PAGE BLANK (08710)